

A Newton's cradle with five silver spheres hanging from thin wires against a light blue background. The spheres are in motion, with some blurred to suggest movement.

Cambridge
Elements



Philosophy of Physics

Why Do We Want a Theory of Quantum Gravity?

Karen Crowther

Cambridge Elements

Elements in the Philosophy of Physics

edited by

James Owen Weatherall

University of California, Irvine

WHY DO WE WANT A THEORY OF QUANTUM GRAVITY?

Karen Crowther

University of Oslo



CAMBRIDGE
UNIVERSITY PRESS



Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781009548083

DOI: [10.1017/9781108878074](https://doi.org/10.1017/9781108878074)

© Karen Crowther 2025

This publication is in copyright. Subject to statutory exception and to the provisions
of relevant collective licensing agreements, no reproduction of any part may take
place without the written permission of Cambridge University Press & Assessment.

When citing this work, please include a reference to the DOI [10.1017/9781108878074](https://doi.org/10.1017/9781108878074)

First published 2025

A catalogue record for this publication is available from the British Library

ISBN 978-1-009-54808-3 Hardback

ISBN 978-1-108-81318-1 Paperback

ISSN 2632-413X (online)

ISSN 2632-4121 (print)

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will
remain, accurate or appropriate.

Why Do We Want a Theory of Quantum Gravity?

Elements in the Philosophy of Physics

DOI: 10.1017/9781108878074
First published online: March 2025

Karen Crowther
University of Oslo

Author for correspondence: Karen Crowther, kcro8477@gmail.com

Abstract: The search for a new scientific theory is typically prompted by an encounter with something in the world that cannot be explained by current theories. This is not the case for the search for a theory of quantum gravity, which has been primarily motivated by theoretical and philosophical concerns. This Element introduces some of the motivations for seeking a theory of quantum gravity, with the aim of instigating a more critical perspective on how they are used in defining and constraining the theory sought. These motivations include unification, incompatibilities between general relativity and quantum field theory, consistency, singularity resolution, and results from black hole thermodynamics.

Keywords: spacetime, theory change, general relativity, quantum field theory, black hole thermodynamics

© Karen Crowther 2025

ISBNs: 9781009548083 (HB), 9781108813181 (PB), 9781108878074 (OC)
ISSNs: 2632-413X (online), 2632-4121 (print)

Contents

1	Introduction	1
2	Incompatibilities between GR and QFT	13
3	The Primary Motivation: Reconciling the Two Pillars	29
4	Unification	35
5	Quantization	39
6	Singularity Resolution	43
7	Black Hole Thermodynamics	55
	References	77

1 Introduction

Usually, when physicists search for a new theory it's because they have encountered something in the world that cannot be explained by the current theories. Such was the case, for example, with the development of quantum mechanics, which began with the anomalous experimental phenomena of the spectrum of blackbody radiation and the photoelectric effect; and also for the development of special relativity, prompted by asymmetries in the predictions of Maxwell's electrodynamics which – famously – did not seem to accord with nature. In these instances, observations precede, and go hand-in-hand along with, the theoretical work: motivating, guiding, and constraining it. Yet, this is not so for the current search for a new fundamental theory of physics, known as *quantum gravity*, which – although potentially related to some empirical problems – has been primarily motivated, guided, and constrained by theoretical and philosophical considerations (but cf. §1.3).

The search for a theory of QG has grown into one of the greatest challenges in modern physics. It is my contention that one of the reasons for the difficulty in finding a theory of QG is that it is not clear what the theory is supposed to achieve, and what, in fact, it *should* be expected to achieve. Thus, a better understanding of the motivations for seeking the new theory can help us to more precisely frame, and eventually answer, the question ‘What is the problem of QG?’ The hope is that critical investigation of the theoretical and conceptual problems which *motivate* the theory can reveal some of the principles (or prejudices) which may underlie these problems. We can then assess whether these principles are well founded, as well as the consequences of adopting them in various roles in theory-development, especially in conjunction with other principles and constraints. It is an aim of this Element to introduce and instigate this new, critical perspective upon the problem of QG. As such, this Element will not cover all of the important developments and philosophically interesting aspects of QG; in particular, discussion of spacetime emergence, spacetime functionalism, and the AdS/CFT duality are conspicuously absent.

So, why do we want a theory of QG? The usual answer refers to ‘two pillars’ of fundamental physics: *general relativity* (GR), providing our best understanding of spacetime, and *quantum mechanics* (QM), our best understanding of matter. Both these frameworks are supposed to be *universal*: unrestricted in their domains of applicability, that is, both are supposed to describe everything in the universe. In practice, however, we typically only *need* to use GR to

¹ *Thermodynamics*, a third pillar of modern physics, is typically neglected in the Primary Motivation for QG, but c.f., §3.4; §7.

describe ‘big stuff’ (the universe at large distance scales), and QM to describe ‘small stuff’ (matter and forces at short distance scales, or, equivalently, high-energy scales). Yet, there are domains of the universe (‘parts of the world’) where both GR and QM are thought to be necessary – where we cannot get away with just using one or the other theory, or any known combination of both. These domains are characterized by extreme densities or temperatures (potentially as high as 10^{93} grams per cubic centimetre, or 10^{32} degrees Celsius), and include the cores of black holes (within the Planck length 10^{-35} m), cosmological singularities such as the ‘big bang’, and the first instants of early universe cosmology. The desire for a description of these domains is part of the primary motivation for seeking the new theory of QG.

In line with my aims of reframing the motivations for seeking QG into positive constraints defining the objectives of the new theory, we might suppose then, ‘that the theory describe these particular domains’ could be the minimal constraint upon the new theory, serving to characterize it. Such a constraint is insufficient, however. GR, on its own terms, *does* describe these domains.² Quantum theory, as well, is supposed to be unrestricted in its domain of applicability. The problem, then, is apparently to find a new theory that describes the domains where both GR and quantum theory are supposed to be necessary, *and* which somehow captures (or ‘takes into account’) the lessons of *both* GR and quantum theory. Let us call this the *Primary Motivation* upon any acceptable theory of QG (although it is very imprecisely defined as stated).

The most straightforward attempts at constructing a theory that fulfils the Primary Motivation – by quantizing GR, treating gravity in the framework of *quantum field theory* (QFT), or otherwise ‘combining’ GR and quantum theory – are beset by various conceptual and technical problems, which have led to their dismissal as unable to fulfil the expectations of a theory of QG. For example, *canonical quantization* of GR leads to the infamous ‘problem of time’ (§2.2); the treatment of gravity in the framework of perturbative QFT is divergent, and thus typically viewed as internally inconsistent (§2.3); while *semiclassical gravity*, which couples classical gravity to quantum matter fields, also leads to divergences, and has been accused of paradox in various thought-experiments (§5). Each of these problems could be referred to as ‘the problem of QG’, since they are obstacles that prevent the acceptance of each approach as a contender for QG. I maintain, however, that these problems are damning only

² On this view, the black hole and cosmological GR singularities are seen as predictions of the theory. This is discussed in §6.

to the extent that they prevent these approaches from satisfying the Primary Motivation on QG. Beyond this, the criticisms of these approaches to QG reveal interesting prejudices we have about the nature of physical theories, and expectations about what QG is to achieve.

The main general motivations for seeking QG that will be explored include:

- *The Primary Motivation*: To have a theory that describes the domains where both GR and QFT are supposed to be necessary, *and* which somehow ‘takes into account’ the lessons of both GR and quantum theory;
- *Incompatibilities between GR and QFT*: To have a theory that provides a coherent picture of the world;
- *Unification*: i. [Minimally] to have GR and quantum theory accounted for by a common framework; ii. [Full] to have a unified theory of all forces, including gravity, as stemming from a single interaction;
- *Putative inconsistency of semiclassical and other ‘hybrid’ approaches to QG*: i. Claims that a theory with quantum matter coupled to a classical spacetime is inconsistent, and thus unable to serve in place of a full theory of QG; ii. Claims that perturbative GR is inconsistent because it is non-renormalizable;
- *Singularity resolution or explanation*: i. To have a theory that addresses the singularity theorems of GR, particularly to describe black holes and the cosmological ‘big bang’ singularity, ii. To have a theory that ‘cures’ the divergences and other mathematical inconsistencies in QFT;
- *Black hole thermodynamics*: To have a theory that describes the evolution of black holes, particularly black hole evaporation as suggested by theoretical work on *black hole thermodynamics*;

However, note that there are (at least) four **other main general motivations for QG** which will not be explored in this Element, namely:

- *Complete cosmology*: i. To have a theory that describes the initial conditions of the universe; ii. To have a theory that solves the cosmological constant problem;
- *Hierarchy problem*: To have an explanation for why the gravitational force is so weak compared to the other fundamental forces.
- *The measurement problem of QM*: To have a more fundamental framework that explains the origin of this problem, and is free of it;
- *The problem of becoming in GR*: To have a more fundamental framework that explains the passage of time (this is briefly mentioned in §2.2).

Each of these represents a number of considerable philosophical challenges in their own right – several have already spawned their own healthily growing philosophical literature. This Element will briefly introduce these related philosophical issues and current debates. Additionally, this Element aims to motivate interest in exploring how these are, could be, and should be translated into positive heuristics to guide theory-development.

1.1 Motivations, Constraints, and Desiderata

The motivations for QG are the reasons for seeking a theory that satisfies the Primary Motivation: answers to the question ‘why do we want this theory in the first place?’ The project that this Element aims to promote starts as a descriptive one, considering *any* motivations cited by physicists, but also looking for other potential motivations. The next stage is an evaluative and prescriptive one. The motivations are framed as principles, and their utility, and implications, assessed in terms of their potential roles in theory-development in various approaches to QG. Both of these stages are merely sketched in this Element, with the aim of inspiring the reader to engage with this project in their own research. Ultimately, the goal is to better understand the theory being sought, and determine whether it is actually what we want or require of QG.

A principle may be used in different ways in different research programmes, and at different stages of theory-development within a given programme (see Crowther (2021) for a discussion of these, with examples from QG). Since I am concerned with motivations here, it is useful to distinguish them from those non-motivational principles and features that share some of the same roles.

Broadly, motivations can be translated into either:

1. *Heuristics or guiding principles*: these are desirable or useful features that can aid in the discovery of the theory by leading to new insights, but which may or may not actually be retained in the resulting theory; they are non-necessary *desiderata*: features that it would be nice if the theory possessed, and which would make us more inclined to accept the theory; or,
2. *(Strong) constraints or criteria of theory acceptance*: the new theory should not be accepted if it is incompatible with the principle (unless there is strong evidence in favour of the theory, and/or the principle is shown to be violated under the relevant conditions); when a motivation is adopted as a constraint, it typically is taken to form part of the definition of the new theory.

Not all of the constraints on QG come from motivations, however, and nor do all desiderata motivate the need for a new theory. Typically, the non-motivational constraints are not taken to define the theory, as the motivational

ones are. Some examples of constraints on QG that are not motivations include those principles coming from current theories within their domains of applicability (e.g., QG predictions should not violate Lorentz invariance in the domains where we know this symmetry holds). There is also the requirement that the new theory *explain* the success of the theories it replaces, in the domains where we know those theories work well. This constraint is related to the Generalised Correspondence Principle, introduced next (§1.2), which also plays a significant role in ensuring the new theory satisfies certain empirical constraints, namely those observations that are explained by current theories. Some of the non-motivational constraints, in spite of being non-motivational, can nevertheless still also be heuristically useful. Some possible examples of desiderata that might not be motivational include the ‘theoretical virtues’ (if these are viewed as guides in developing a new theory rather than motivations to seek a new theory), as well as unexpected explanation of (theoretical or empirical) problems that are not hitherto thought to be quantum-gravitational.

The relationship between motivations, constraints (necessary features), and desiderata (non-necessary features), is illustrated in Figure 1.

Motivations, constraints, and desiderata can all play important roles as *means of confirmation* or as indicators of *pursuit-worthiness* of a program: compatibility with the principle (or empirical result) serving to increase credence in the theory, or as being suggestive of the theory’s potential future success, if it were to be developed (‘pursued’). I believe this holds regardless of whether

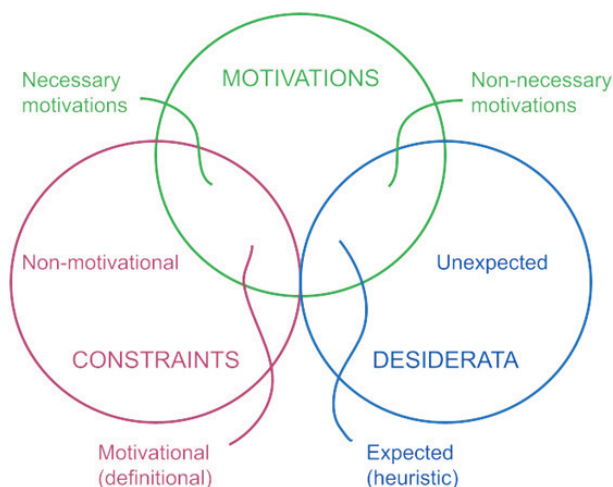


Figure 1 *Motivations, constraints, and desiderata*: Motivations may be necessary (constraints), or non-necessary (desiderata).

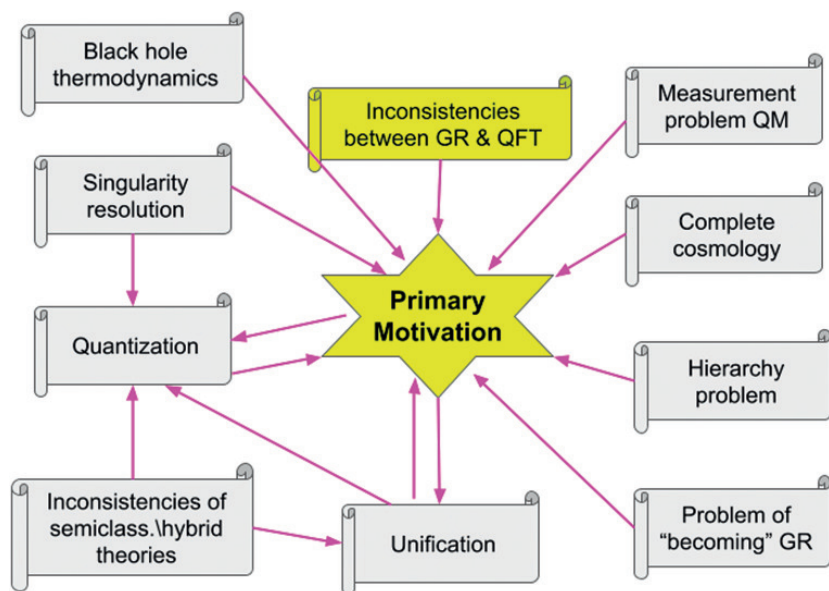


Figure 2 Various motivations as related to the Primary Motivation. Arrows represent ‘motivates’; unification and quantization are possible routes towards fulfilling the Primary Motivation, so can also be thought of as motivated by the Primary Motivation (as well as independently motivated).

the principle has been used in the development of the theory, or whether the theory’s possession of the feature in question is unexpected (‘prediction’).³ Keep in mind, too, that principles are not always purely theoretical, but can be based on empirical results. Empirical results can serve in various roles, though different results may feature in different roles, as discussed further (§1.3).

So, how are the motivations listed earlier used, and how are they related to the Primary Motivation? These questions will be explored in the rest of this Element, though my conclusions here are certainly more provisional than conclusive, and I invite further engagement. The tentative list below is illustrated in Figure 2.

- Incompatibilities between GR and QFT motivate the Primary Motivation, representing the constraint of *consistency* (which itself requires philosophical assessment in this context), but also frustrate attempts at satisfying it;
- Unification motivates the Primary Motivation, but also represents different (i–ii) ways of satisfying the Primary Motivation; it is not a constraint, but a desideratum;

³ Some philosophers accuse this of illegitimate ‘double counting’, but I’ve never understood this argument. If the theory satisfies various constraints and desiderata (including the theoretical virtues), then this is evidence in its favour; though, it can perhaps be argued that the surprise factor adds more weight.

- *Putative inconsistency of semiclassical and other ‘hybrid’ approaches to QG* arise from attempts to satisfy the Primary Motivation, underpinned by the constraint of consistency; they are not motivations for the Primary Motivation, but are motivations for particular ways of satisfying it (through unification, typically taken to indicate the need for quantization), which I argue are not necessary ways;
- *Singularity resolution or explanation* motivates the Primary Motivation, through the constraint of consistency; can serve as a guiding principle and desideratum, but not all singularities in GR and QFT require resolution by QG;
- *Black hole thermodynamics* stems from attempts to bring together GR and QFT (plus thermodynamics), it motivates the Primary Motivation; leads to various desiderata and potential constraints;
- *Complete cosmology*: these are problems related to empirical observations which QG is expected to solve, and so they motivate the Primary Motivation; arguably, they represent constraints on QG;
- *Hierarchy problem*: stems from empirical concerns; motivates the Primary Motivation, but is (arguably) a desideratum rather than a constraint;
- *The measurement problem of QM*: as above;
- *The problem of becoming in GR*: as above.

1.2 Relative Fundamentality

The general motivations for QG listed earlier may be seen as reasons for not treating GR and QFT as fundamental. There are different notions of, or requirements for, fundamentality in physics,⁴ but the key notion here concerns *completeness* (or *maximality*) of description. In this Element, I take it that QG need not be a fundamental theory, in the sense of being a *final theory* or *theory of everything*. All that is required is fulfilment of the Primary Motivation. As such, we should not require as strong constraints on the QG all those features that are expected of a final theory.

However, QG is expected to be more fundamental than GR and QFT. The notion of relative fundamentality here can be understood simply as an asymmetric relation of non-causal dependence (Crowther, 2018). QG is more fundamental than GR and QFT by virtue of its having a broader domain of applicability, one that subsumes the domains of success of the previous theories – QG describes ‘more stuff’ than either of these frameworks. GR and

⁴ Morganti (2020); Crowther (2019).

QFT are thus expected to be *effective* theories (or frameworks): valid as approximations to QG in the restricted domains where they are known to be successful.

Asymmetric dependence between an older, ‘predecessor’ theory O , and a newer, ‘successor’ theory, N , may be demonstrated through the establishment of particular *correspondence relations* between them, such that O is taken to be *derivable* from N . There are various types of inter-theory correspondence relations,⁵ though the most familiar are characteristic limiting relations (a standard example is the transition from relativistic to classical mechanics in the *Newtonian limit* of small velocities compared to the speed of light, that is, $v/c \rightarrow 0$). The idea is essentially what Nickles (1973) calls ‘reduction₂’, where N is said to ‘reduce to’ or *recover* O in the domains where O is known to be successful. This is standardly taken as a constraint on any new theory of physics that is more fundamental than its predecessor. This is also a way of satisfying the *Generalised Correspondence Principle*, which forms an important constraint on QG, and states that any two theories whose domains of applicability overlap share approximately the same results within these domains.⁶

Given that QG is taken to be more fundamental than GR and QFT, we expect that GR and QFT be derivable from QG in this way. (In spite of this, the *recovery* of QFT has not been appealed to as a constraint on QG to the same extent as has the recovery of GR or spacetime, except within the context of string theory.) On the other hand, QG is not standardly expected to be more fundamental than the framework of quantum theory: most approaches to QG attempt to satisfy the Primary Motivation by simply providing a quantum description of spacetime (or, more precisely, the quantum physics underlying classical spacetime). According to this ‘standard perspective’, QG is itself a quantum theory, rather than a more fundamental framework supposed to explain quantum mechanics (§3). This said, however, it is a significant, and underexplored, question as to whether the explanation of quantum mechanics *should* be understood as part of the motivation for QG – for instance, if we expect QG to explain the measurement problem of QM. There are some approaches that do take this as a goal, most notably the work by Penrose (2002, 2014). If standard quantum theory is modified in the construction of QG, or QG were to represent a new

⁵ Radder (1991), identifies three types of correspondence relations, and Hartmann (2002), describes seven.

⁶ Here, I follow Butterfield & Isham (2001, p. 79) in taking ‘results’ to include theoretical propositions as well as observational ones, and even ‘larger structures’ such as derivations and explanations. The Generalised Correspondence Principle is discussed by Post (1971), but the statement here comes from Crowther (2020).

framework, of which quantum mechanics were a special case, then QG would be considered more fundamental than quantum theory.

1.3 Empirical Constraints

Although QG is primarily motivated by non-empirical concerns, the need to be consistent with empirical observations is still a principle that is used extensively in the search for QG. Three different types are employed: empirical observations that *are not* explained by current theories; empirical observations that *are* explained by current theories; future empirical observations. I briefly describe each in turn.

1.3.1 Empirical Results Not Explained by Current Theories

These are motivations, but may or may not be constraints. Both GR and quantum theory are remarkably successful in describing the world: there are scant observations that they leave unexplained. Nevertheless, there are some open problems related to observations. The two biggest ones are the problems of dark matter and dark energy. There are various astrophysical observations that cannot be explained by GR given the known amount, and type, of matter in the universe – these can be explained by positing an abundant amount of a new form of matter, known as *dark matter*, with behaviour very different from known matter. Additionally, the expansion of the universe is observed (via, e.g., measurements of supernovas) to be accelerating in a way that cannot be accounted for according to the main currently accepted cosmological model based on GR and the known matter in the universe; this underlies a problem known as *dark energy*.

Although the problems of dark matter and dark energy may be related to QG – and to each other – they are not standardly treated as such. The observations associated with dark matter, for instance, are instead generally taken to motivate searches for new particles beyond the Standard Model – though other candidates, such as primordial black holes, do not require new physics. Additionally, most approaches to QG have not taken these anomalies as empirical evidence motivating the need for a more-fundamental theory of QG. There are some exceptions to this ‘traditional’ position, however – and, indeed, there are a growing number of papers related to QG that treat the solution to these problems as a guiding principle, means of confirmation, or indicator of pursuit-worthiness for a theory of QG.⁷ This raises interesting questions as to how, or whether, the solutions to the problems of dark energy and dark matter should

⁷ The most prominent is Verlinde (2017), but see also Calmet & Latosh (1998); Kastner & Kauffman (2018), Oriti & Pang (2021).

be used in the search for QG. Perhaps these empirical anomalies *should* motivate the search for a more fundamental theory of matter and spacetime, and be counted as part of ‘the problem of QG’?

1.3.2 Empirical Results Explained by Current Theories

These are not motivations, but are constraints. It is a criterion of acceptance that QG reproduce the empirical results explained by current theories, in accordance with the Generalised Correspondence Principle (§1.2). The Generalised Correspondence Principle plays many roles in theory development and assessment: Its wide use is reflective of its status as a ‘shortcut’ to empirical results, through the demonstration the new theory is consistent with the empirical results already covered by the older theory. In QG, the Generalised Correspondence Principle takes a more specific form, as the requirement that GR be appropriately *derivable* from QG in the domain where GR is known to be successful. This correspondence between QG and GR has overwhelmingly dominated as a constraint of interest – Carlip (2001, p. 927), for instance, refers to this as the ‘zeroth test’ of QG.

1.3.3 Novel Empirical Results

These are motivations and constraints. Novel empirical results come from *QG phenomenology*: a unique field of research that aims to reciprocally connect QG to observable phenomena, by building models that bridge the considerable gap (many orders of magnitude) between them. So far, there has not been any philosophical work dedicated to exploring QG phenomenology in general, in spite of the importance of the field, and its offering intriguing connections to philosophy of scientific modelling and experiment.⁸ QG phenomenology has led to numerous results, including tight empirical constraints on any possible violation of Lorentz invariance (this is a key symmetry of GR that some approaches to QG suggest may be broken under some circumstances, but QG phenomenology has revealed no evidence of its being violated).

While QG phenomenology typically connects with cosmological and astrophysical observations, it can also potentially connect with laboratory experiments, such as ‘tabletop’ Gravitationally Induced Entanglement experiments, which may provide a ‘witness’ of the underlying quantum nature of gravity in the non-relativistic limit, using superpositions of Planck-mass bodies.⁹

⁸ The reader may take this as a hint of an interesting research project to explore.

⁹ These include Bose et al. (2017); Marletto & Vedral (2017); philosophical exploration in Adlam (2022); Huggett et al. (2023).

Other novel empirical tests of QG phenomena may (arguably) come from *analogue experiments*, which are experiments whose Source system (the particular system being experimented on) is thought to be analogous to the Target system (the general class of systems that the Source system is supposed to represent) *and* where the Target system itself is inaccessible (the inaccessibility of the Target system is part of the definition of analogue experiments, following (Crowther et al., 2021)). An example is the case of acoustic ‘dumb hole’ models, in which sound waves in fluids encounter a horizon, and which are described by equations that are isomorphic to those of black holes. These were originally proposed by Unruh (1981), using *analogical reasoning*, since the dumb holes are supposed to be analogous to black holes. There has been interest in whether or not analogue models of black holes in the laboratory are able to confirm the existence of *Hawking radiation* – which, as discussed in §7, would give us insight into properties of black hole spacetimes that are potentially interesting for QG.¹⁰

1.4 Approaches to QG

There are various different research programs, or *approaches*, towards QG, with different motivations, assumptions, formalism, and of varying degrees of completeness. Each of these have their own problems, which, from their own perspective could be called ‘the problem of QG’. This Element, however, is written from a general perspective, discussing issues relevant to *any* possible theory of QG: indeed, it aims, specifically, at better understanding the common features by virtue of which we can even conceptualize these different approaches to QG as being ‘approaches to QG’. The idea, as stated, is to clarify, and begin to assess, the most basic motivations for seeking QG in the first place. The different approaches to QG each assign varying levels of importance to the various principles and motivations, and utilize them in different ways. As I go through the motivations and principles, I will mention the ways they have been used in some approaches to QG as illustration. But, given the limited space of this Element, I will not introduce or explain what these approaches are in any detail.

We can broadly classify the approaches of three types, owing to their historical development as stemming from the different methods of attempting to quantize gravity (see, e.g., Carlip, 2001; Rovelli, 2000). As described in §2.3, one natural approach to quantizing GR is to use familiar techniques from

¹⁰ Steinhauer (2016) has reported to have detected Hawking radiation using Bose-Einstein condensates in the laboratory. See also Barceló et al. (2011); Weinfurter et al. (2013). Philosophical discussion: Crowther et al. (2021); Dardashti et al. (2017, 2019).

QFT to treat gravity, which broadly captures the ‘particle physics perspective’ (or ‘high-energy physics perspective’) towards QG. An early attempt utilizes quantum perturbation theory, which is known as the perturbative *covariant quantization* approach to QG. This approach to quantization splits the GR metric into a background part plus small perturbations; the perturbations are quantized, while the background remains classical and non-dynamical. [Splitting apart the dual role played by the metric field in GR, and restoring spacetime as a fixed, static, ‘background’, goes against the principle of *background independence* in QG (§2.1), and was thus objectionable from the ‘GR perspective’ on QG]. This approach was shown to suffer from gravity being non-renormalizable (see §2.3). In response, the years 1980–2000 saw the development of research into *non-perturbative* QG, and the majority of research since has focused on this (Loll et al., 2022). The main approach to QG that historically stems from the covariant quantization methods is *string theory* (see, e.g., Zwiebach, 2009).

Asymptotic safety is an approach based on the Wilsonian idea of renormalization, and the conjecture that quantized gravity is non-perturbatively renormalizable (discussed in §2.3).

Closely related to the covariant approach is *Feynman quantization*, which applies Feynman’s path integral (sum over histories) quantization techniques to the GR metric. Some important current examples of approaches that aim to implement these methods include *causal dynamical triangulations* (Loll, 2019), and *causal set theory* (Surya, 2019).

Alternatively, the *canonical* or *Hamiltonian* quantization of GR aims to construct QG as a theory of the fluctuations of the metric as a whole, so no fixed metric is involved. This approach represents the ‘GR perspective’ on QG. It involves first casting GR into canonical form, and then using canonical quantization methods. This requires splitting spacetime into space and time, with a fixed topological structure. The metric information is represented, in full, by operators on a Hilbert space. The goal is then to find the eigenvalues of the operators along with their transition probabilities. While the canonical QG approaches aim at preserving the principle of background independence, they suffer more manifestly from the *problem of time*, §2.2 (though problems of time also arise in other approaches to QG). The main approach to QG stemming from the canonical quantization procedure is *loop quantum gravity* (although now loop quantum gravity also is being developed using a covariant formalism, see Rovelli (2004); Rovelli & Vidotto (2014)).

Two other classes of approaches mentioned in this Element are *hybrid theories* and *emergent gravity* approaches (§3.3,5). Note, however, that this general classification by no means exhausts the range of approaches pursued in QG.

2 Incompatibilities between GR and QFT

QM is not a theory, but a framework within which particular theories are formulated; similarly, QFT is also a framework – one which combines QM and SR. Basically, it extends QM to fields, that is, systems with an infinite number of degrees of freedom, defined on a fixed, flat, continuum spacetime of SR. The framework uses spacetime, but it does not *describe* spacetime. Theories formulated in this framework are said to be *background dependent*, because the spacetime serves as a ‘background’ in the theory (see §2.1). The Standard Model of particle physics is a theory formulated in the framework which we might call ‘conventional QFT’ (CQFT) Wallace (2006, 2011). The framework treats all matter as composed of particles, which are understood as local excitations of quantum fields; the fundamental forces are themselves represented by quantum fields, whose corresponding excitations interact locally with the other particles, depending on their type. Any dynamical field, according to QFT, is quantized. Incorporating gravity into this framework would entail treating gravity as a field whose force is mediated by a particle called the graviton.

GR is a theory *of* spacetime; the equivalence of gravitational and inertial mass means we can understand gravity as a property of spacetime itself, rather than a field propagating on a fixed spacetime background. It stands in contrast to QFT for several reasons. Most basically, GR says that spacetime is a dynamical field rather than a fixed, static background structure. It is a *nonlinear* theory (while QM is a linear theory), in that spacetime ‘reacts’ to matter and energy, and in turn, the behaviour of matter and energy are affected. To use the famous 12-word summary of Wheeler, ‘Spacetime tells matter how to move; matter tells spacetime how to curve’ (Wheeler, 2000, p. 235).¹¹ But, from the perspective of GR, we could say that gravity is not really a force at all—it does not ‘act’ on objects (Maudlin, 2012). Objects simply exist in spacetime, and GR tells us that an object’s inertial path is determined by the curvature of spacetime (rather than, e.g., conceiving of gravity as a field that deflects objects from their inertial paths).¹²

The formalism of GR can be defined by the Einstein-Hilbert action,¹³

$$S_{EH} = \frac{c^4}{16\pi G} \int_{\mathcal{M}} d^4x \sqrt{-g} (R - 2\Lambda) - \frac{c^4}{8\pi G} \int_{\partial\mathcal{M}} d^3x \sqrt{h} K \quad (1)$$

where g is the determinant of the metric, R is the Ricci scalar, and Λ is the cosmological constant. The first integral is over a spacetime region \mathcal{M}

¹¹ Cf. (Misner et al., 2017, p. 5).

¹² For other formal compatibility issues, see §3.6.3 of Rickles (2006).

¹³ I’m here following the presentation in Kiefer (2007a).

(a four-dimensional manifold of spacetime points, encoding the topology and differentiable structure), and the second integral is defined on the boundary, $\partial\mathcal{M}$, of this region. This term is required for a consistent variational principle; here, h is the determinant of the three-dimensional metric, and K is the trace of the second fundamental form. We also consider a ‘matter action’, S_m , for non-gravitational fields, which give rise to the energy-momentum tensor,

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S_m}{\delta g^{\mu\nu}} \quad (2)$$

which acts as a ‘source’ of the gravitational field. From the variation of $S_{EH} + S_m$, we obtain the Einstein Field Equations (EFE),

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \quad (3)$$

where $G_{\mu\nu}$ is the Einstein tensor describing the curvature of spacetime, $g_{\mu\nu}$ the Lorentzian metric tensor, encoding the geometry, $R_{\mu\nu}$ is the Ricci curvature tensor. A model (i.e., a spacetime) of GR is specified as $M = \langle \mathcal{M}, g_{\mu\nu}, T \rangle$ where the two tensors $g_{\mu\nu}$ and $T_{\mu\nu}$ satisfy the EFE (3).

So, on the one hand, we have GR describing classical geometry as a dynamical field coupled to classical matter, and on the other hand, we have QFT which uses a fixed, non-dynamical classical geometry, and which describes quantum matter, saying that all dynamical fields should be quantized. How can they be combined? One straightforward option would be to modify (3) to take into account the quantum nature of matter, by replacing $T_{\mu\nu}$ with the expectation value of the *quantum energy-momentum tensor operator*, $\langle \hat{T}_{\mu\nu} \rangle$, to obtain,

$$G_{\mu\nu} = \frac{8\pi G}{c^4} \langle \hat{T}_{\mu\nu} \rangle \quad (4)$$

These are the *semiclassical Einstein equations*, where gravity stays classical while the other fields are quantum. Note that the quantum energy-momentum tensor operator is not only difficult to compute mathematically, but its expectation value is also difficult to understand physically.¹⁴ Nevertheless, it is an operator that acts on states $|\psi\rangle$ of a material quantum system, and thus obeys the Schrödinger equation, with Hamiltonians describing both the dynamics of matter with itself, and with gravity,

$$i\partial_t|\psi\rangle = \hat{H}_{\text{matter} + \text{gravity}}|\psi\rangle \quad (5)$$

A system described by equations (4–5) is referred to as *semiclassical gravity*. While the semiclassical Einstein equations may be of value as approximations,

¹⁴ Cf. Huggett et al. (2023); Kiefer (2007a); Wald (1994).

leading to insights into the low-energy regime of QG, they face several serious conceptual and theoretical difficulties if treated as exact equations at the fundamental level of QG.¹⁵

More generally, the basic intuition that semiclassical gravity is inconsistent can be appreciated by analogy with the case of electromagnetism: Bohr & Rosenfeld (1933) analysed an equation akin to the ‘semiclassical’ equation for electromagnetism and demonstrated that the electromagnetic field had to be quantized in order to be consistent with the quantized matter it couples to.¹⁶ The basic idea is usually taken as the uncertainty relations in the quantized system spread to (‘infect’) the coupled non-quantized system. In the case of semiclassical gravity, the uncertainty in the position of a quantized gravitating object would lead to quantum uncertainty in the gravitational field, so the gravitational field itself should be quantized. Although such arguments are heuristically very powerful, the consensus is they cannot by themselves compel the quantization of gravity (see §5).

Nevertheless, this does suggest another natural way to combine GR and quantum theory: attempt to quantize the gravitational field, similar to the way in which the electromagnetic field was quantized. This approach, too, faces extraordinary conceptual and theoretical difficulties. The quantization of the electromagnetic field in quantum electrodynamics (QED) results in a theory of quantum fluctuations of the electromagnetic field against a well-defined classical spacetime background. The attempt to quantize gravity, however, means subjecting some of the properties of spacetime to quantum fluctuations. We thus run into trouble in giving a mathematical characterization of the quantization procedure itself without a well-defined background spacetime. Even if we are able to achieve this, we then face trouble in giving a physical account of the theory that results – for example, a fluctuating metric would seem to imply a fluctuating causal structure and spatiotemporal ordering of events, which makes it difficult to define equal-time commutation relations in the quantum theory.

One of the most prominent clash of principles between GR and QM has to do with the superposition principle of QM. It seems, firstly, that superpositions conflict with general relativity’s fixed causal structures – it is difficult to see how we could understand superposed causal structures of GR. Possibly relevant for this, however, are recent ‘quantum switch’ experiments, which show

¹⁵ While Großardt (2021, 2022) discusses ‘three little paradoxes’ of semiclassical gravity, Erik Curiel describes a ‘panopticon of problems’ with semiclassical gravity and black hole thermodynamics; see his talks linked at <http://strangebeautiful.com/phil-physics.html>. See also Kiefer (2007a).

¹⁶ Butterfield & Isham (2001, §3.1.2.), Huggett & Callender (2001).

quantum superposition of causal structure (Chiribella, 2020; Goswami et al., 2020). Furthermore, as (Penrose, 2004, §30.10–30.11) argues, using a thought experiment called ‘Schrödinger’s lump’, there may be a conflict between the superposition principle of QM and the principle of general covariance of GR.¹⁷ Penrose (2014) uses this as part of an argument for the ‘gravitization of quantum theory’, as opposed to the quantization of gravity.

The superposition principle has also, however, been used to argue for the quantization of gravity. This argument comes from Feynman, who considers a Stern-Gerlach type experiment in which a spin-1/2 particle is guided to two counters. The counters are connected to an indicator which is either up when the particle arrives at counter 1 or down when it arrives at counter 2. The indicator itself is a ball of macroscopic dimensions (1 cm), which would then be in a superposition of being in two positions. Since the ball is macroscopic, its gravitational field would also be in a superposition. ‘We would then use that gravitational field to move another ball, and amplify that, and use the connections to the second ball as the measuring equipment. We would then have to analyse through the channel provided by the gravitational field itself via the quantum mechanical amplitudes. Therefore, there must be an amplitude for the gravitational field.’ (Feynman, quoted in Zeh (2011), p. 66).¹⁸ The idea is that the gravitational field itself must be described by quantum states subject to the superposition principle. As Kiefer (2013), states, this is not an argument that demonstrates the necessity of quantizing gravity, but rather ‘is an argument based on conservative heuristic ideas that proceed from the extrapolation of established and empirically confirmed concepts (here, the superposition principle) beyond their present range of application. It is in this way that physics usually evolves’ (p. 2). Kiefer (2013) points out that while there are several such heuristic arguments, there are none that *logically demand* the quantization of gravity.

In the rest of this section, I discuss three other notorious issues that arise related to the attempt to quantize GR: the problem of *background independence*, §2.1, the *problem of time*, §2.2, and the *perturbative non-renormalizability* of quantum GR, together with the principle of *UV-completeness* §2.3, §6.1.

2.1 Background Independence

Background independence, and the principle of *relationalism* are used widely in the search for QG. Several authors view the history of physics in terms of

¹⁷ Cf. Penrose (2014); Weinstein & Rickles (2021).

¹⁸ The original report is from 1957 and republished in Dewitt & Rickles (2011).

the increasing relationalism of its theories, and thus motivate the quest for QG along these lines. Smolin (2020), for instance, adopts relationalism as a methodological imperative to ‘make our theories more relational by eliminating background structure’.¹⁹ One might wonder, however, whether it actually represents a *motivation* for QG, rather than just being, for example, a constraint or desideratum upon a new theory of QG. On the one hand, we might argue that the goal of having a background independent theory is not a reason to seek QG, since we already have a theory of gravity that is background independent, namely GR. But, while background independence can certainly be viewed as part of the motivation that drove the development of GR, it is not straightforward to identify what exactly is meant by background independence in GR as I discuss below. On the other hand, though, QM and QFT are straightforwardly not background independent theories, so a drive for increasing relationalism through the elimination of background structure could promote bringing these theories in closer connection to GR, and thus to motivate the Primary Motivation.

There are four general, but distinct, principles associated with background independence:

- a. That the theory not feature an absolute, fixed, ‘background’ structure (typically: spacetime);
- b. That the theory be compatible with *relationalism*;
- c. That the theory be *generally covariant*;
- d. That the theory be *non-perturbative*.

The first of these (a.) is most properly called ‘background independence’: basically, a theory that is background independent does not posit the existence of a fixed background object, or require the existence of such an object in order to define the properties of other objects in the theory. A well-known example of a background *dependent* theory is Newtonian mechanics, which features absolute space and time as fixed background structures used to define position and motion. For Newton, space was not only *absolute* and *immutable* (meaning that it is fixed, existing unchanged throughout all time, totally unaffected by matter, but also responsible for affecting matter, being used to define the properties of – or itself conferring properties upon – matter) but also *substantial* (meaning that space itself is a ‘substance’, as real as matter, and which can exist independently of matter).

¹⁹ Others who motivate QG along these lines include Rovelli (2004), and proponents of the ‘discrete approaches’ to QG, for example, Caravelli & Markopoulou (2011); Ambjørn et al. (2006); Loll (1998); Markopoulou (2009a); Williams (2006).

For QG, background independence standardly means that the theory not feature a fixed (background) spacetime, but there are several meanings, corresponding to different notions of ‘fixity’ (Butterfield & Isham, 1999, §4.1), as well as the different (layers of) structures that could be identified as ‘spacetime’ (metric, manifold, affine structure, etc.). There are (at least) four proposals that attempt to formally capture the idea of background independence in the sense of (a.), earlier. Note that each is a separate proposal:

- i. A theory is background independent if it features no absolute objects (i.e., structures that are the same in all models of the theory);
- ii. A theory is background independent if it has no formulation which features fixed fields (i.e., there is no formulation of the theory that is not diffeomorphism invariant);
- iii. A theory is background independent if its solution space is determined by a generally-covariant action (all of whose dependent variables are subject to Hamilton’s principle, and represent physical fields);
- iv. A theory is background independent if all of its physical degrees of freedom correspond to geometrical degrees of freedom.

For details on (i–iii), Pooley (2017), and for (iv), Belot (2011) (a comprehensive discussion of background independence in classical and quantum gravity can be found in Read (2023)). Each of these attempted characterizations has its own difficulties, however – and, arguably, none are adequate to do the job.²⁰

The idea of (b.) relationalism in regard to background independence comes from the thought that a theory that does not rely on a fixed background structure will have to define the properties of its objects purely in terms of the relations between these objects, rather than relying on the background structure. In this sense, relationalism is opposed to absolutism (that there be a structure that affects objects without itself being affected), as well as substantivalism (that space itself has an independent existence, rather than just being a means of representing the relations between objects). Also, an absolute object is typically understood as being fundamental, in the sense that it not ontologically dependent on any further, more basic, underlying objects. So, when it is claimed that a background independent theory is relational, it means that the properties in the theory are defined only with respect to one another, and that these relationships are not fixed (in the sense of being non-dynamical), but evolve according to the equations of motion of the theory. While we might for this reason identify ‘relationalism’ and ‘background independence’ as synonymous,²¹ Rickles

²⁰ Further details: Giulini (2007); Pitts (2006); Pooley (2017); Read (2023).

²¹ As does Smolin (2006, p. 204).

(2008) argues against the identification on the grounds that a substantival view of spacetime is also compatible with background independence.

A major driving force for seeking a background independent theory is the desire to preserve, or continue, what is seen as one of the most important insights of GR. Yet, while GR is taken as the prime exemplar of a relational or background independent theory, it is difficult to precisely identify what it is that ‘makes GR special’ in this regard. The intuitive notion we have is that GR is special because – owing to the theory’s *diffeomorphism invariance* or *general covariance* – we cannot define the physical observables of the theory without solving the dynamics, or, in other words, that ‘there is no kinematics independent of dynamics’ (Stachel, 2006). However, it is not easy to define explicitly the idea of a dynamical object along the lines of ‘being solved for’. The main problem, following Kretschmann (1917), is that we can turn *any* background field into a dynamical object by making it satisfy some equations of motion, however physically vacuous they might happen to be. Even the metric in special relativity is able to be made dynamical in the sense that it satisfies some equations of motion – hence special relativity can be made background independent, which then conflicts with our basic intuitions of what background independence is supposed to be (Rickles, 2008).²²

It is not impossible that general covariance (c.) be used in defining ‘what is special’ about GR,²³ and thus feature in a principle of theory selection, though general covariance can not *by itself* be the principle of background independence. However, most of the recent literature has instead moved on in to principle (a.) in their attempts to salvage the general intuition that GR is special by virtue of possessing some particular conception of background independence. In this context, the proposal (i) earlier fails for GR (and any theory with massive fields),²⁴ and while proposal (ii) arguably works for GR, it is difficult to implement as a guiding principle in the search for a new theory (Read, 2023). Nevertheless, the success of GR as a theory of spacetime, together with the significant heuristic role of the principle in its discovery, mean, however, that background independence (in some sense, which may, arguably, be any of (ii–iv)) continues to be viewed as a desirable principle in QG, in all roles.

Background independence is not standardly considered a *motivation* for QG: it is not a reason for seeking a new theory in the first place. One way of framing

²² There is a long-running debate on the question of ‘what makes GR special’, for example, Earman (2006); Pooley (2010).

²³ Cf. Norton (2003); also Brown (2005, §5.3.1).

²⁴ See fn. 20.

it as a motivation would be to argue (1) QM/QFT are not fundamental because they are not background independent, and (2) that fundamental physics, as it progresses through theory-change, trends towards increasingly relational, or less-background-dependent theories. Such an argument by induction is not very strong, however, without additional reasons for expecting a fundamental theory should be background independent, or an argument as to why background independent theories are better or more successful than background dependent ones. The main argument is that background independent theories are *more explanatory* than background dependent ones because they do not require structures to be fixed in advance, that is, put in ‘by hand’, but instead have their structures follow from the dynamical laws and principles specified by the theory itself. In this case, *if* background independence is a *motivation* for QG, then it would mean that QG is motivated by the need to explain the background elements of QM/QFT (and potentially also those in GR), and itself to not feature these elements. If we were to reject the principle as a motivation for QG, then, arguably, the principle need not be treated as a *constraint* on QG, unless QG were conceived of as a final theory – if the theory required background structures, not themselves explained by the theory, then we would be motivated to search for a more fundamental theory that did explain these structures.

Finally, there is the distinct notion, (d.), that the theory be *non-perturbative*. Perturbative treatments of gravity involve fixing a background metric (e.g., Minkowski metric) and studying gravitons propagating on this background. The desire that a theory be non-perturbative is a natural one, given that perturbation theory is a set of techniques used to construct approximations when the exact theory is unknown or unsolvable. These techniques are of limited applicability, and can easily lead to problems when misapplied. Some problems with perturbative gravity are considered in §2.3.

I close this section with a brief mention of background independence in string theory and some other approaches to QG. From early on, string theory has been intensely criticized for being a background *dependent* approach – on account of its being based on perturbative methods, and describing strings moving on a background spacetime – it apparently fails on all (a.–d.).²⁵ In recent years, however, the criticism has died down, as string theorists came to appreciate the significance of background independence, and the theory itself shows hints of being background independent in various ways. (For this reason, the philosophical debate concerning the importance of background independence in QG is no longer so active, having achieved its goal, in a sense.) As Read (2023, §5.2) demonstrates (building on a claim made by Huggett & Vistarini (2015)), there

²⁵ See, e.g., Smolin (2006).

are in fact several senses in which string theory (considered both at the level of spacetime fields, and at the level of worldsheet fields) may be considered background independent, on some of the definitions (ii–iv) earlier, given that the background fields in the theory are required to be dynamically coupled together in the same dynamical equations of GR (the Einstein field equations, plus higher order corrections), and thus are not ‘fixed’ after all, but possess background independence in ways similar to GR.

Additionally, there are attempts to construct a non-perturbative version of string theory; the most celebrated development being the *AdS/CFT duality* (also known as the *Maldacena conjecture*, after Maldacena (1998)). This is a relationship – a conjectured exact equivalence – between a string theory featuring gravity, describing closed strings propagating on a spacetime (anti-de Sitter space (AdS)), known as the ‘bulk’, and a gauge theory without gravity (a conformal field theory (CFT)), defined on the boundary that contains the bulk spacetime. In the regime where the string theory is *strongly coupled*, and computations difficult – the CFT is *weakly coupled*, meaning that perturbative techniques can be used to make calculations. Since the theories are supposed to be equivalent (describe the same physics), the weak coupling regime of the CFT can be used to shed light on the strong coupling regime of the string theory.²⁶ This is closely related to the holographic principle, §7.4. While string theorists hope to ultimately arrive at a theory that is background independent, the approach began as a background dependent one (perturbative string theory). Credence in the approach increases by its being shown to be background independent, it may also feature as a criterion of theory-acceptance.

These comments also apply, broadly, to the program of *asymptotic safety*, which is an approach to QG that treats gravity in the framework of QFT, but avoids the problems associated with perturbative non-renormalizability (see §2.3). While the approach is premised on non-perturbative QFT techniques, it still relies on splitting the full metric $g_{\mu\nu}$ into a fixed, but arbitrary, background metric $\bar{g}_{\mu\nu}$ plus a dynamical fluctuation field $h_{\mu\nu}$. The principles of diffeomorphism invariance and background dependence (in the sense of (a.)) pose challenges – both technical and conceptual – for the theory, associated with its use of the *functional renormalization group* techniques and the *background field method* (Bonanno et al., 2020).

Other approaches to QG have placed heavy emphasis on background independence from their inception, including LQG (Rovelli & Vidotto, 2014), causal set theory (Surya, 2019), causal dynamical triangulations (Ambjørn et al., 2009), group field theory (Oriti, 2009), and other ‘pre-geometric’

²⁶ See, e.g., Polchinski (2017).

approaches (Caravelli & Markopoulou, 2011). These approaches strive for a fully background independent theory, and sell themselves as upholding what they see as a central insight of GR (and of any acceptable theory). These approaches argue that the key to finding a successful theory of QG is in complete (as much as possible) elimination of any background structure or geometry. Adopting Smolin's (2006) slogan, the idea is 'the more relational the better'. These approaches use the principle of background independence in all roles.

The principle of background independence is now widely accepted as an important feature for QG to possess: it is presented as the main candidate for satisfying the condition that QG 'take into account' the lessons of GR in the Primary Motivation (§3.3). Nevertheless, it is an open question whether background independence is a sufficient means for satisfying this aspect of the Primary Motivation: there may be other principles from GR that could serve this role. Background independence is furthermore desirable since it functions as means of principle correspondence, connecting QG with GR, and thus helping, in part, to satisfy the Generalised Correspondence Principle and establish a connection between QG and known empirical results.

2.2 Problem of Time

There is an incompatibility between the treatment of time quantum theory and in GR. According to quantum theory, time is external to the system being studied: it is *fixed* background in the sense that it is specified from the outset and is the same in all models of the theory. In GR, however, time forms part of what is being described by the theory. Time in GR is subject to dynamical evolution, and it is not 'given once and for all' in the sense that it is the same across all models (Butterfield & Isham, 1999). This clash between the external time in quantum theory and the dynamical time in GR leads to problems in combining the theories, and is often taken as a suggestion – along with other heuristics, such as the suggestion of 'quantum foam', or the Planck length as a minimal length – that time 'disappears', or 'does not exist' in QG.

The *problem of time* is a cluster of problems, usually framed along the lines just suggested, but specifically as manifesting in the context of *canonical QG*. This is a class of approaches to QG that attempts to quantize the canonical (or Hamiltonian) formulation of GR, via familiar canonical quantization techniques. The oldest and most straightforward of such approaches is known as *quantum geometrodynamics*, which proceeds in a spirit similar to Schrödinger's original heuristic approach to quantization (via the Hamiltonian-Jacobi formalism). In this context, one finds the primary form of the 'problem of

time': the supposed 'dynamical' equation of the resulting quantum theory (the Wheeler-de Witt equation) does not feature a time variable, and so presents us with a 'frozen formalism'.²⁷ Yet, as Thébault (2021) emphasizes, the problem as I have just set it up, is misleading in the sense that it does not in fact arise from 'from forcing a background independent theory of spacetime onto the Procrustean bed of quantization with respect to a background time'. Instead, Thébault (2021) argues, 'the problem arises generically from the manipulation of a particular class of classical theories, that includes general relativity, according to the standard formal steps that are preparatory for quantization. That is, the problem of time becomes apparent in the process of preparing any reparametrization invariant theory for quantization' (p. 390).

Thus, the problem of time (or, at least, the *global* problem of time that Thébault (2021) focuses on) may not be specific to GR, and may arise even before quantization. This does not ameliorate the fact, however, that there are numerous problems related to the treatment of time in QG. The relevant question for this Element is whether these problems represent *motivations* for seeking the new theory, or rather challenges that crop up once we embark on this quest. As mentioned in §1, one motivation for seeking QG (which is not discussed in this Element) is the 'problem of becoming': if we are dissatisfied with the lack of 'real change', or temporal passage, in the standard 'block universe' picture of spacetime as presented by SR and GR, a possible option is to see if we can find a more hospitable notion of time in the more-fundamental theory beyond GR. This type of motivation is based on a perceived inconsistency between time in SR/GR, and our own human experience of time (i.e., of time flowing, and the experience of events occurring 'Now'). One *prima facie* difficulty with this strategy is, of course, analogous to that faced by attempts to overcome the determinism of GR (i.e., in the hopes of finding 'free will') by looking to more fundamental theory: given that relativity does describe our macroscopic world, and that any more-fundamental theory needs to be compatible with this in the regimes where GR is successful, we may question what relevance any concept of time in QG has to our everyday experiences of 'change' and 'becoming'. Plausibly, though, the connection with becoming could be relevant in the context of QG cosmology, where the theory is expected to replace the Standard Model of cosmology (based on GR), including the idea of the 'big bang', in describing the initial conditions of the universe.

More generally, we might take the disparate treatments of time in GR and QM as motivation for seeking a new theory that has a single, coherent account

²⁷ It is important to emphasize that there are several difficulties for interpreting the Wheeler-de Witt equation. Detailed reviews of the problem of time: Isham (1993); Kuchař (1999). Recent philosophical treatments: Anderson (2017); Thébault (2021).

of time. The exploration of time in QG has spawned a huge amount of discussion, with all sorts of different solutions being proposed: from those that feature space without time (e.g., Barbour (2000); Gomes (2020)), time without space (e.g., Markopoulou (2009b); Smolin (2013, 2020), neither time nor space (e.g., Rovelli, 2004), and disagreements over the implications of QG for the notion of ‘becoming’ (e.g., Dowker, 2020; Rovelli, 2020). So, is ‘the problem of QG’ about finding a fundamental description of time, or of timeless fundamental physics that can nevertheless explain the description of spacetime in GR, or of providing an account of change and/or becoming? Arguably, none of these are necessitated by the Primary Motivation. In order to translate these into positive constraints or heuristics for the new theory, we need to evaluate the arguments (including metaphysical arguments) for one solution over another – assuming, of course, we follow Huggett & Wüthrich (2013), in not outright rejecting such theories as unphysical or ‘empirically incoherent’ on account of not featuring spacetime fundamentally.

2.3 Perturbative Non-renormalizability of Quantum GR

A natural way of approaching QG is to utilize familiar techniques from QFT – for example, using quantum perturbation theory to treat gravity. The framework of QFT within which the Standard Model of particle physics is formulated relies upon perturbative techniques in order to make predictions in most cases. A perturbative calculation of any particular physical process involves a summation over all possible intermediate states, and this is done at all orders of perturbation theory (though, in practice, often only the first few terms are taken, and it is hoped that the higher-orders decay rapidly). The dynamics of QFT is local, utilizing point sources and interactions – it features arbitrarily small distances and, thus, arbitrarily large momenta. This, as well as the integration over all the momentum-energy states, implies that there are an infinite number of intermediate states. Consequently, perturbative calculations within the theory lead to divergent integrals (if the terms are not sufficiently suppressed). These infinities are dealt with in particular QFTs by the process of *renormalization*, which first imposes a finite-valued cutoff in momenta (so that the integral is not taken over all possible high values of momenta), and then the parameters of the theory are re-defined in terms of ‘physical’ coupling constants that absorb the divergences, and whose values are then fixed by experiment. A theory is *renormalizable* if only a finite number of counterterms are required to absorb the divergences, whereas a *non-renormalizable* theory requires an infinite number of such terms. Because an infinite number of terms cannot be determined experimentally, a non-renormalizable theory is not predictive.²⁸

²⁸ See, e.g., Weinberg (1995). In the context of QG, see (Kiefer, 2007a, §2.2.2).

The strong and electroweak interactions are renormalizable, and thus the Standard Model of particle physics is a renormalizable theory. Gravity, however, is non-renormalizable: the Einstein-Hilbert action (1), treated in the perturbative framework, has divergences appear in loop diagrams at first order (in the matter case, or second order in the matter-free case), and there is an expectation that infinitely more infinities appear at higher orders.²⁹ The divergences occur at high energy scales, near the Planck scale, and we are prevented from calculating anything in this regime.

In standard QFT, however, even non-renormalizable theories can be predictive (and impressively so), if they are treated as *effective field theories*. This is thanks to the development of *renormalization group* (RG) techniques, which led to a framework for studying the way in which the physical coupling constants change depending on the energy scale at which the theory is being applied (the couplings are said to ‘run’ or ‘flow’ with the RG).³⁰ Effective field theories are valid at low energies compared to some high energy valued cutoff scale: they are not formally predictive to arbitrarily high energies; that is, they are not thought to be *UV-complete* (§6.1).

Perturbative quantum GR can be treated as an effective field theory which makes predictions for low energies compared to the Planck scale. It has proven to be an extremely good approximation for all the present experimental tests of gravity (Burgess, 2004; Wallace, 2022). However, the low-energy nature of its predictions is taken to imply that quantum effects are important where gravitational fields are very strong, such as inside black holes or cosmological singularities. In other words, its failure here is taken to indicate the domains of necessity of QG. It also means that this approach towards finding QG fails to satisfy the Primary Motivation, and thus motivates the search for a new theory.

However, it is possible that the proliferation of infinities in the ultraviolet (UV) regime does not signal the failure of GR treated in the framework of QFT, but instead signals the limitations of the perturbative approach in this regime. The *asymptotic safety scenario* claims that the physical couplings, in the *non-perturbative* RG flow, do not actually diverge, but instead flow to a finite value: a ‘fixed point’ in the UV.³¹ This is similar to QCD, where the couplings flow to a fixed value of 0 in the UV, and the theory is said to be asymptotically free

²⁹ The one-loop divergences were shown by 't Hooft & Veltman (1974); cf. Bern (2002).

³⁰ The RG and associated philosophy of effective field theory is rich in philosophical interest; see, e.g., Batterman (2018); Butterfield & Bouatta (2015); Cao & Schweber (1993); Crowther (2015); Williams (2021), and Kaplan (2005); Manohar (1997); Weinberg (2009) for the physics.

³¹ This was proposed by Weinberg (1979); cf. Donoghue (2020), Eichhorn (2019), Niedermaier & Reuter (2006).

(the theory is non-interacting in this regime). In the case of gravity, however, the fixed point is supposed to be non-zero (the theory is interacting in at least one of the couplings), and the theory is said to be *asymptotically safe*, since the physical quantities are ‘safe’ from divergences as the cutoff is removed (taken to infinity). If there is a fixed point, then following the RG trajectory (almost) to it, one can in principle extract unambiguous answers for physical quantities on all energy scales (Niedermaier & Reuter 2006). At the fixed point, the dependence on the UV cutoff is lost, and the theory is *scale invariant*: it does not change as smaller length scales are probed. While this approach may satisfy the Primary Motivation for QG, it does suffer other problems, including a conflict with the principle of unitarity. This *problem of unitarity*, means – roughly – that probabilities will not be conserved in time.³²

2.4 UV Completeness

The failure of perturbative quantum GR to provide a predictive theory at the Planck scale is motivation for the Primary Motivation. The problem, from this perspective, is that the theory fails in the regime where we expect quantum effects to be significant for gravity. Nevertheless, the problem is often misdiagnosed as the failure of the theory to be UV-complete, and as indicating the need for a UV-completion of this theory. Thus, UV-completeness is often taken as a criterion of theory acceptance for QG. Here, following Crowther & Linnemann (2019), I argue that it need not be.

A theory is UV-complete if it is formally predictive to all possible high energy scales (or, equivalently, all short distance scales). In other words, the theory can be used to generate predictions at all possible short distance scales. Being UV-complete does not guarantee that all of these predictions are *correct*, however (Newtonian mechanics is UV complete, for instance, since it does not formally ‘break down’ at any short distance scale, and yet we know that it is not the correct theory to use at all short distance scales). There are several ways of obtaining a UV-complete theory, including having a QFT that is non-perturbatively renormalizable, such as quantum chromodynamics.³³ Yet, the idea of UV-completion is *not restricted to QFT*, and there are other ways in which a theory may be UV-complete, apart from being renormalizable.³⁴

³² Another approach, known as *conformal gravity*, which is perturbatively renormalizable, also suffers this problem of unitarity (Gomes, 2020, p. 179).

³³ *Perturbative* renormalizability is insufficient to establish UV-completion, since a perturbatively renormalizable theory may still face a Landau Pole, as is the case in quantum electrodynamics §6.2.

³⁴ Butterfield & Bouatta (2015); Crowther & Linnemann (2019); Dvali et al. (2011); Zee (2010).

In the heyday of QFT, UV-completion was usefully employed as a criterion of theory success in QFT: that is, a means of selecting theories that were not only viable, but *physical*. The development of the RG and effective field theory techniques (and associated philosophy) led to a change in this perspective, where we recognize that theories can be physical without being fundamental – effective field theories correctly describe the world within their limited domains, and are consistent when applied within these domains. Nevertheless, it seems there remains a sort of hangover, where the general perception of UV-incomplete theories as being faulty, or mathematically inconsistent, still lingers.

Although approaches such as asymptotic safety are guided by the principle of UV-completion, QG need not be a UV-complete theory, *unless QG is conceived of as a fundamental (final) theory*. If we conceive of QG in a more minimal sense, then all that is required to satisfy the Primary Motivation is a ‘UV-better’ theory, predictive at the Planck scale (Crowther & Linnemann, 2019). Following Figure 3, we may thus distinguish between two options: either the ‘amalgamation’ of QG (as not ToE)³⁵ plus the Standard Model (i.e., a non-unified set of theories) are together fundamental (i), or they are not. If they are not, then (some of) these theories are emergent, and they can either emerge from a ToE (ii), or they emerge from another amalgam (iii). And so on. Ultimately, there are three possible scenarios: (1) some non-unified set of theories is fundamental, (2) there is a ToE, or (3) there is no fundamental level, but a never-ending tower of theories.

It is only in scenario (i) in which QG is necessarily UV-complete. This scenario is distasteful, however, in that it represents a disunified amalgamation of several theories as standing in place of a final ToE. The disunified nature of the amalgam would prevent us from believing it fundamental,³⁶ and we would push on in the search for a more unified theory beyond. This would be the case in spite of all the theories in the amalgam being UV-complete. The lesson from this is that UV-completion should only be used as a constraint on QG if you accept that, at the fundamental level, there is no final theory, but rather a disunified collection of several theories.³⁷

³⁵ If one includes the criterion that QG be a ToE, then QG is the ToE in (ii), rather than part of the amalgam (i).

³⁶ Other concerns, such as arbitrariness of the values of the parameters within each theory and the lack of explanatory power, would also contribute to this thought (Crowther, 2019).

³⁷ In some approaches to QG, UV-completion may conflict with other principles that are desirable and otherwise apparently viable, aside from unification. For instance, if UV-completion is obtained by cutoff, it may pose problems for Lorentz invariance, and, in the ‘higher-derivative approaches’ and asymptotic safety scenario, UV-completion apparently conflicts with the principle of unitarity.

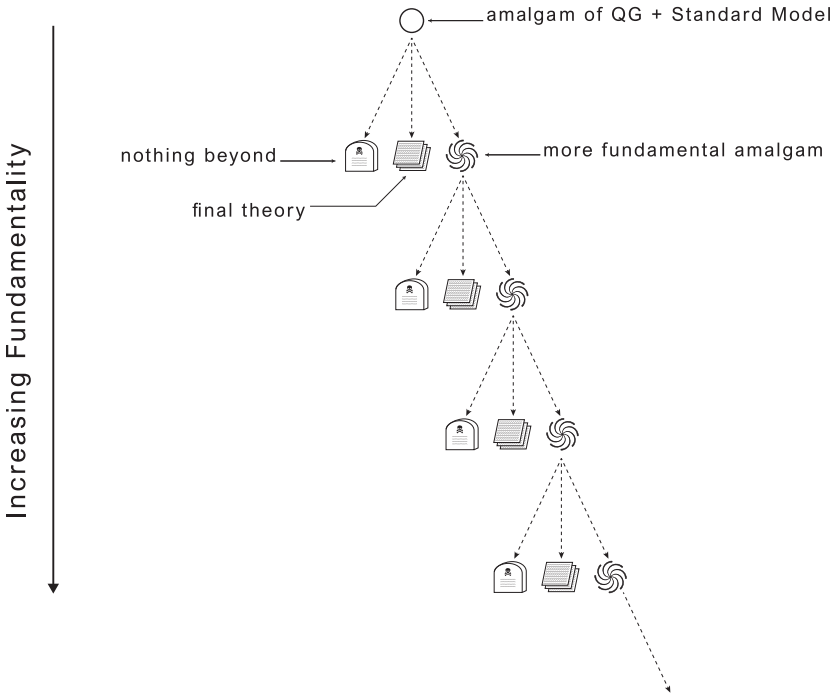


Figure 3 *Options on the road towards higher-energy physics:* Suppose we have a non-unified set of theories ‘amalgam’ of QG and the Standard Model.

It is possible that there is no deeper theory beyond, or that this amalgam emerges from a unified ‘final theory’, or that it emerges from another amalgam. This could happen at each ‘level’ of inquiry as we seek increasingly fundamental theories.

String theory is supposed to be a ToE, rather than just QG – as such, it must be UV-complete. The extendedness of the theory’s basic entities is very likely to ensure that it is UV-complete, although no proof has yet been found that it is (see Hagar (2014, §7.2), Dawid (2013, Ch. 1)). Textbooks on string theory often present an analogy between the perturbative non-renormalizability of quantum GR and that of 4-Fermi theory, which is a non-renormalizable theory that was revealed to be the low-energy limit of the *renormalizable* theory of electroweak interactions. Similarly, proponents of string theory claim that string theory is the renormalizable theory underlying perturbative quantum GR. Thus, the alleged UV-completeness of string theory is treated as a means of confirmation or indicator of pursuit-worthiness.

Other approaches to QG use UV-completion in different roles, particularly as a guiding principle, constraint, and means of confirmation. In some cases, including in LQG, UV-completion is obtained automatically in virtue of the theory naturally describing a minimal length scale (Rovelli, 2004). Here, without

taking UV-completeness as a constraint on QG, it may make sense to treat the principle as a means of confirmation, especially given the consilience of the idea of minimal length with other motivations for QG, such as singularity resolution, §6.

3 The Primary Motivation: Reconciling the Two Pillars

I’ve characterized the Primary Motivation for QG very roughly as a theory that describes the domains where both GR and QFT are supposed to be *necessary*, and which somehow ‘takes into account’ the lessons of both GR and quantum theory. Three points require more elaboration: ‘describes’, ‘domains’, and ‘taking into account’.

3.1 ‘Domains’

Smeenk (2013) describes the ‘overlapping domain’ argument for QG, arguing that the early universe falls within the domains of applicability of both QFT and GR, and thus requires a theory of QG for its description. As stated earlier, however, both GR and quantum theory are supposed to be universal theories, so technically their domains of applicability overlap completely! Thus, QG should also be a universal theory, not restricted to particular scales or parts of the world. Properly, the motivation for QG is to describe domains where QG effects cannot be neglected – where the predictions of QG are thought to differ from those of GR and QFT. In other words, it should describe the overlapping domains of *necessity* – rather than applicability – of GR and quantum theory, as viewed from the current paradigm.³⁸

Wallace (2022), as well as the physicists interested in QG phenomenology, emphasizes that not all the domains of QG are exotic and inaccessible. Also, in fact, we have a working theory that satisfies the second condition of the Primary Motivation as a combination of GR and quantum theory: this is just QFT applied to GR,³⁹ *and* which describes QG phenomena and observations – that is, observations of QG phenomena that we in fact already have access to – *and* the theory is relatively well-confirmed (depending on the different domains where it is applied, its support ranges from strong, to moderate, to weak). This is a *low energy theory of QG* (LEQG); it is an effective field theory, meaning that it is valid in a particular, restricted domain. This

³⁸ Of course, though, from the perspective of the imagined future paradigm, QG is supposed to replace GR and QFT in these domains, so – strictly – the presently incumbent theories are not supposed to be necessary, nor even correct, in these domains.

³⁹ This is not just semiclassical gravity (eq. 4), but also a QFT treatment of fluctuations using the *background field method*, such that GR is evaluated via the path integral formalism of QFT (Wallace, 2022).

domain includes QG phenomena – being self-gravitating objects that also require QM to describe – in the semiclassical regime, both non-relativistic (describing planets and white dwarfs), and relativistic (describing stars, neutron stars, supernovas, and cosmology), and also the perturbative regime (having some support from observations of the solar system and cosmic microwave background).

The domain of LEQG, however, is low-energy⁴⁰ – the theory is not thought to describe the Planck scale (explained next), the interior of black holes, or very early universe cosmology. It is these domains for which we need a full theory of QG, and which properly serve to characterize the Primary Motivation for QG. Thus, ‘the problem of QG’ from the perspective of high-energy physics, is that LEQG does not describe the high-energy (UV) regime. (The theory also faces particular theoretical and conceptual problems, but these are not what prevents it from satisfying the Primary Motivation.)

The Planck scale is characterized by combining the fundamental constants from our current theories: the speed of light, c , the gravitational constant (Newton’s constant), G , and the quantum of action (the reduced Planck’s constant), \hbar . It was shown by Planck in 1899 that there is a unique way (apart from numerical factors) to do so. We arrive at the Planck length, l_P , the Planck time, t_P and the Planck mass (equivalently, the Planck energy), m_P , respectively.⁴¹

$$l_P = \sqrt{\frac{\hbar G}{c^3}} \approx 1.62 \times 10^{-35} \text{m} \quad (6)$$

$$t_P = \sqrt{\frac{\hbar G}{c^5}} \approx 5.40 \times 10^{-44} \text{s} \quad (7)$$

$$m_P = \sqrt{\frac{\hbar c}{G}} \approx 1.22 \times 10^{19} \text{GeV} \quad (8)$$

In cosmology, we also refer to the Planck temperature T_P , and Planck density, ρ_P ,

$$T_P = \frac{m_P c^2}{k_B} \approx 1.41 \times 10^{-32} \text{K} \quad (9)$$

$$\rho_P = \frac{m_P}{l_P^3} \approx 5 \times 10^{93} \text{g/cm}^3 \quad (10)$$

where k_B is Boltzmann’s constant.

These are obtained by dimensional analysis, a familiar and useful technique in physics. Theories are typically characterized by their own fundamental

⁴⁰ ‘Low-energy’ here refers to anything below the Planck energy scale, however, so it includes all known physics!

⁴¹ See, e.g., Kiefer (2007a), §1.1.3.

constants, and because this dimensional analysis combines both the characteristic constants of GR and QM in a unique way to form a new set of characteristic constants, it is standardly thought that these characterize the regime where QG is necessary.

In spite of the popularity of the Planck scale as characterizing the domains of QG, the dimensional analysis that gives us the Planck scale is heuristic: it does not establish that this is the characteristic scale of QG. The relevance of the Planck scale can be criticized as a naive general estimate. Nevertheless, it is possible to make the case that the Planck scale is the relevant scale at which to expect new physics, based on arguments from both particle- and gravitational-physics which may be taken to suggest that our current picture of familiar ‘large scale’ (i.e., energies lower than the Planck energy) physics is complete (Held, 2019). Such a conclusion, however, neglects the problems of dark matter and dark energy (the cosmological constant Λ), which may be considered as part of QG phenomenology.

Perhaps stronger than the heuristic arguments for the relevance of the Planck scale is the argument that the Planck scale is where LEQG – as standardly conceived, perturbatively – breaks down, so necessarily a new theory of QG is required here (§2.3).

The other domains that characterize QG are the interior of black holes and the very early universe. This is motivated by the idea that GR ‘breaks down’ in the vicinity of curvature singularities because of the neglect of expected quantum effects in these domains (§6). Another motivation comes from the theoretical predictions of black hole thermodynamics (§7).

3.2 ‘Describes’

Ideally, for a theory to *describe* particular domains, the theory gives predictions in these domains that are tested and not falsified. This is a challenge given the domains in question, however. More minimally, then, we might require that the theory give predictions that are physically reasonable when applied in these domains – although, what ‘physically reasonable’ means is a non-trivial question in this context! More work needs to be done to explore this, particularly in connection to discussions in general philosophy of science.

Of course, a theory being formally predictive, and that the predictions are taken as ‘physically reasonable’, in these domains does not guarantee that what it says is correct when applied here, but we can usefully appeal to the empirical constraints described in §1.3, including correspondence relations with GR and QFT to provide at least some guidance.

3.3 'Takes into Account Both QM and GR'

Finally, consider the second main condition of the Primary Motivation: that the theory *take into account both QM and GR*. This condition is motivated by the success of GR and QM, and also the fact that both are expected to be necessary in the particular domains indicated earlier – and yet, each are thought to be inaccurate here. There are several ways of understanding this condition, representing different viewpoints on the problem of QG. Ashtekar expresses some of these,

Everything in our past experience tells us that the two descriptions of Nature we currently use must be approximations, special cases which arise as suitable limits of a single, universal theory. That theory must be based on a synthesis of the basic principles of general relativity and quantum mechanics. This would be the quantum theory of gravity that we are seeking. (Ashtekar, 1995, p. 186) (p. 186)⁴²

This quote demonstrates two ways in which QG could take into account GR and quantum theory: by recovering both theories in their respective domains of applicability (through correspondence relations, including limiting relations), and combining the principles of both theories. These two conditions represent what we might call the 'new framework perspective' of the problem of QG. It is also described in Ashtekar & Geroch (1974), in a passage where the authors state that GR and QM each can be understood as a distinct 'body of universal rules', which means that the quantization of gravity is essentially different from quantizing the electromagnetic field: 'From this viewpoint, then, the problem is to obtain a new body of rules which suitably encompasses the essential features of those of quantum mechanics and of general relativity. (p. 1214).

This 'new framework perspective' on the problem of QG, however, is not the standard one adopted by physicists working on QG.⁴³ Rather than seeking a new theory that replaces GR and QM, and recovers both theories as approximations in their respective domains, most physicists tend to treat QM as fundamental and GR as not.⁴⁴ For instance, Loll et al. (2022, p. 16) expresses the sentiment that the 'universality of quantum theory' is one of the fundamental principles of modern physics: the physical world at the fundamental level is governed by quantum laws, and that the classical picture is only an approximation, valid at sufficiently low energies and on sufficiently large scales.

⁴² Cf. Ashtekar (2005).

⁴³ Note that the labels 'new framework perspective' and 'standard perspective' are my inventions, and that Ashtekar's work is ambiguous between the two views.

⁴⁴ Which is not to say that GR fails to embody some principles, for example, background independence, that are thought to be fundamental.

Thus, we find at least two different ways of satisfying the condition that QG ‘take into account’ the lessons of GR and QM. What I have called the *New Framework Perspective* holds (1) we need a *new body of rules* that express insights of both GR and QM, and (2) we need to recover *both* GR and QM as approximations in their respective domains. The *Standard Perspective*, however, holds (1) we need a *quantum theory*, and (2) from this we can recover GR as an approximation in the domains where we know GR is successful (on such a view, the recovery of QM is not required). While I will not here argue for one view or the other, I would at least like to suggest that the fundamentality of QM is an assumption that has not been questioned seriously enough in the search for QG, and that the problem of QG may well be that we require a whole new framework that could ‘explain’ the success of quantum theory just as it would explain the success of GR.

There are more-specific ways of satisfying the second main condition of the Primary Motivation. Butterfield & Isham (2001), §3.1.3. describes four different types of approach to QG: (1) quantize GR, (2) GR as the low-energy limit of a quantization of a different classical theory, (3) GR as the low-energy limit of a theory that is not a quantization of a classical theory, (4) start ab initio with a radically new theory. Approaches 1 and 2 represent the Standard Perspective, approach 4 represents the New Framework Perspective, while approach 3 could represent either perspective, depending on whether or not the theory is formulated in the framework of quantum theory as we currently understand it. Any of these approaches could produce a theory that accords with the condition of ‘taking into account’ both GR and QM, if it satisfies the two necessary aspects described earlier. Nevertheless, the by-far dominant approach is (1), where not only is QG supposed to be a quantum theory, but one in which gravity is quantized.

It is possible that QG take into account QM without itself being a (purely) quantum theory. The most familiar way of thinking about this is along the lines of semiclassical gravity (§2) or other ‘hybrid’ approaches (§5), which attempt to couple the classical and the quantum as a sort of ‘amalgamation’ (as opposed to a unification) or ‘mongrel gravity’ theory (a term borrowed from Mattingly (2009); cf. Tilloy (2018) for discussion of some other hybrid theories). The New Framework Perspective countenances the possibility that the framework of QM – as it stands – is not applicable in QG, and gets modified somehow (in which case QG would be more fundamental than both GR and QM, §1.2, but need not be a full unification of GR and QM §4).

As stated earlier, Penrose believes that the measurement problem of QM points to the need for QG, and that the unitarity of QM will be modified, so that QM will become a non-linear theory at the Planck length. More recently,

the unitarity of QM has been called into question in the context of quantum cosmology, where Cotler & Strominger (2022) replace linear unitary time evolution with linear isometric time evolution (note: this is very different from the quite radical proposal of Penrose). Other recent work in the New Framework perspective is the hybrid, ‘post-quantum classical gravity’ theory which modifies both GR and QM Oppenheim (2023). This approach is motivated by problems with semiclassical gravity as well as perceived inadequacies to mainstream responses to the black hole information loss paradox.

QG can ‘take into account’ the lessons of GR and QG by means of *principle correspondence*: the new theory features one or more principles from each of the older theories (Hartmann, 2002). As part of taking into account GR, it is popular to utilize background independence as a principle-correspondence, featuring in both GR and QG (§2.1). For principle correspondence in QM, the favourite principle to use is quantization (§5), which accords with the Standard Perspective. These principles are naturally, and conservatively, motivated by the universality of both GR and QM.

3.4 Three Pillars Perspective

An increasingly significant perspective upon the problem of QG is that it involves bringing together not just GR and QM, but also the third pillar of modern physics: thermodynamics (also including statistical mechanics). There are several approaches to QG that tie gravity to the thermodynamic concept of entropy. For instance, Jacobson (1995) argues that the Einstein Field Equations can be derived from the proportionality of entropy and black hole horizon area together with the first law of thermodynamics. Padmanabhan (2004, 2010) and Verlinde (2011, 2017) also present arguments for gravity being an emergent phenomenon of entropic origin. According to the ‘emergent gravity’ approaches, discussed in §5, spacetime is an effective thermodynamic entity; as such, a quantization of gravity (the metric field) would not lead us to the ‘micro’ degrees of freedom which we seek to describe by a more fundamental theory of QG. (These micro degrees of freedom, however, could themselves be described by a quantum theory, and so these approaches need not represent the New Framework perspective.) The link between GR, QFT and thermodynamics is motivated, and further revealed, through the theoretical results of black hole thermodynamics, discussed in §7.

3.5 Delimitations

The Primary Motivation is supposed to represent the minimal definition of QG: any approach that does not (attempt to) satisfy it would not be accepted, *by the*

current mainstream consensus, as an approach to QG. The definition thus serves to delimit what would ‘count’ as an acceptable theory, and necessarily excludes alternative possibilities. We should bear in mind that there may be other solutions to the problems discussed here, which are overshadowed if we take them exclusively as problems that motivate QG (as a theory that satisfies the Primary Motivation). Examining the particular motivations in more detail can help us also to explore the other possibilities for their resolution, or reconsider the need to resolve them at all.

This is nicely illustrated by the example of the differing attitudes one can take towards spacetime singularities (§6): not all cases of singularities motivate the need for a new theory, and in the cases where the (particular) singularities do motivate a new theory, it is not certain that these point to the need for QG rather than a different theory formulated ‘at the level of current theory’ (i.e., a new classical theory of gravity such as a modification of GR). There is the possibility that other motivations for QG may also be addressed by reconsidering our current best theories, including their principles and formulation, rather than moving to a theory of QG.

4 Unification

Unification is a traditional guiding principle in physics, and is often viewed as means of producing successful theories. Familiar examples (representing various different ideas, and degrees, of unification) include Maxwell’s theory of electromagnetism, which unified light as well as the electric and magnetic forces; the electroweak theory, which unified the electromagnetic force and the weak force; and even GR, with its identification of inertial mass with gravitational mass, and spacetime with gravity. There is a tendency to view the history of physics as a history of unification, and the path forward as one of continuing this trajectory to its ultimate end in a final, unified theory: Salimkhani (2018) calls this the ‘physicist’s tale’, following Maudlin (1996), who states that it ‘has become so pervasive as to rank almost as dogma’. It is illustrated, for instance, in Figure 4.

As shown in Figure 4, unification is a way of motivating a theory of QG, here illustrated as a final ToE (but see Figure 3 for an alternative picture, illustrating QG as part of a non-unified fundamental level). For those inclined towards unification, the current situation in physics – the split picture of the world it presents – is unsettling, and calls us to question the fundamental nature of both GR and the framework of quantum theory and the Standard Model of particle physics. This may be the case *even if we accept* that there is a usable theory that treats gravity in the same framework as the (other) fundamental forces, namely

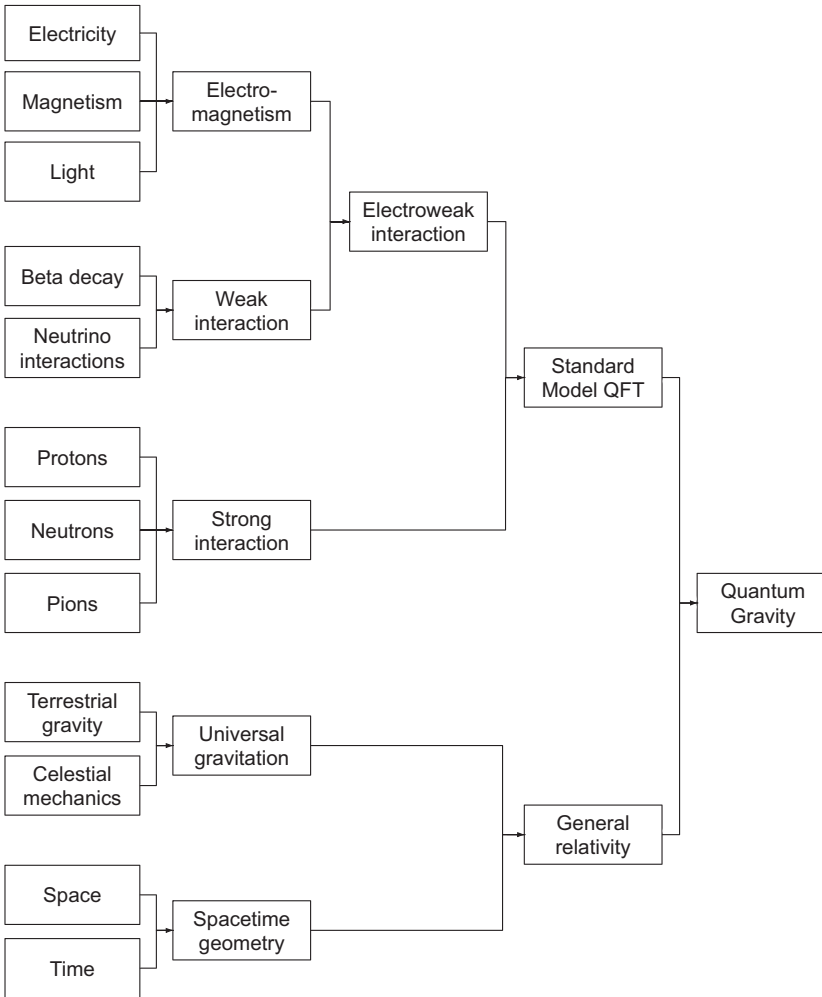


Figure 4 The ‘physicist’s tale’ of unification.

LEQG (introduced in §3). In other words, the problem is not that we cannot combine GR and QM in a single theory: the problem, on this view, is that the theory is not properly *unified*.

What is meant by unification? Maudlin (1996) argues that there are several degrees (or levels) of unification, which exist between a ‘lower bound’ that falls short of unification, and an ‘upper bound’ that represents *perfect unification*. The lower threshold states that unification is not merely that two (or more) theories are consistent with one another, or that they share a common dynamics, or that there exists a law-like connection (nomic correlation) between physical forces. The upper bound of perfect unification requires that there be a single

theory that describes all phenomena as *the same* – as fundamentally stemming from a single origin, for example, as manifestations of a single entity or interaction (‘All is One’). And, ‘it is this deeper sense of unification, the idea that all the physical forces are at base one and the same, which contemporary physicists invoke when they speculate on the theories to come’ (Maudlin, 1996, p. 132). Morrison (2000) refers to this type of unification – where two phenomena hitherto thought to be distinct are identified – as *reductive* unification.⁴⁵

Unification is typically seen as an external motivation for QG (the desire to unify two separate frameworks), but it can also be considered as an internal motivation (within a single framework), if the lack of unification *within* the Standard Model of particle physics is taken as a problem motivating QG. Although the Standard Model can be written as a single theory, it appears as a disjointed amalgam of separate (particle) fields, rather than a unified theory in Maudlin’s sense, and this drives many physicists to seek a more unified theory beyond. But why unification? Unification is generally regarded as an epistemic virtue, conferring support for a theory, and being used as a means of justification of a theory (note that this can be the case even without requiring or implying the metaphysical assumption that the world itself is unified); several authors have given Bayesian analyses of unification as an epistemic virtue (Myrvold, 2003). Unified theories have also been argued as being more explanatory, testable, falsifiable, and successful than non-unified (or less-unified) ones (Schindler, 2018, §1.1.4). (Additionally, unification can be used *heuristically* in guiding theory-development, as described by Kao (2019)).

String theory is often promoted as being a unified theory, as well as being a ToE, since it treats gravity as on par with the fundamental forces of the Standard Model, and all stem from the same basic physics (the behaviour of strings). Yet, QG is not necessarily a unified theory, nor a ToE. In the first case, it may be a semiclassical or hybrid theory, which is a non-unified combination of GR and QFT, and in the second case, it may just be a quantum theory of gravity, and not a theory that combines gravity with the Standard Model forces. Requiring QG to be a unified ToE is only justified if we take QG as a *final* theory (i.e., ‘nothing beyond’ in the schema illustrated in Figure 3).

There may be reasons for not wanting to unify gravity with the forces of the Standard Model of QFT, however, depending on how one interprets GR. The typical interpretation of GR is a geometrical one, according to which gravity is not properly a force at all, but the curvature of spacetime. As Maudlin

⁴⁵ It is just one of thirteen different forms of unification described in Morrison (2000).

(1996, p. 133) states, ‘Objects do not couple to the gravitational field, they merely exist in space-time.’⁴⁶

This interpretation of GR stands in contrast to the ‘particle physics perspective’ which drives string theory. Instead, it – along with the principle of background independence (§2.1) – motivates approaches to QG which primarily aim at a consistent quantum description of (more fundamental physics responsible for the low-energy existence of) spacetime, such as LQG and causal set theory. In such approaches the principle of background independence is prioritized over the principle of unification.

But in general, perfect unification of GR and QM is not a standard motivation for QG. If it were, this would mean that both GR and quantum theory (as a framework) are not fundamental, but must somehow be recovered from QG – using appropriate limiting procedures and approximation techniques – in the regimes where they are known to be successful. If any approach to QG could not demonstrate that this can be achieved, then its acceptance would be unlikely. But, this is not what most approaches aim at; they aim to retain the standard framework of quantum theory, rather than reduce it to a more basic theory. Arguably, even unification in a weaker sense is not properly called the problem of QG, if we would be satisfied with a ‘hybrid’ approach to QG, or asymptotic safety.

Salimkhani (2018) adopts the particle physics perspective and argues that the problem of QG is not the need to unify – nor even to combine – GR and QM. He uses results by Weinberg to argue that GR can be reduced to (derived from) the combination of SR and QM (i.e., QFT), and that we have LEQG; thus that the problem is that we need to find the correct theory at high energy scales. More generally, Salimkhani (2021) argues that unification need not be understood as a goal of physics at all. Unification is considered an external influence on physics – *external* here meaning that it is driven by philosophical assumptions, such as metaphysical, metatheoretical, or epistemological considerations that are imposed on physics. Against this, Salimkhani argues that, instead, unification naturally arises in physics as a consequence of the more basic (or genuine) aims and methods of physics, that is, factors properly *internal* to physics itself, such as empirical adequacy and theoretical consistency.

While the primary motivation for QG is to find a theory that describes the domains where both GR and QFT are supposed to be necessary, and which somehow combines GR and QM, it seems we should not understand the problem – as it currently stands – as to find a unified theory. This is not to say, however, that the goal of unification may not be heuristically useful, or

⁴⁶ Cf. Maudlin (2012).

that unification could not be an epistemic (or explanatory) virtue, or a means of confirmation for a theory of QG. These are questions that are left open, and depend on how else we conceive of the problem of QG.

5 Quantization

Quantization is the procedure by which a quantum theory is produced from a classical one. There are three main ways in which we can attempt to perform a quantization of gravity, and each of these faces its own challenges (§1.4). However, when we refer to the need for QG to be a quantum theory, or more specifically, a theory of quantized gravity, this does not imply that it must be a theory obtained through the quantization of GR. Recall the four different ways of constructing QG identified by Butterfield & Isham (2001), described earlier (§3), wherein the quantization of GR was only one option.⁴⁷ In other words, even on the Standard Perspective, we can have a theory of QG that is quantum, but not produced through the quantization of GR. And, further, if we adopt the recondite New Framework perspective, we may deny not only that QG need be obtained through a quantization of GR, but also that QG need even be a quantum theory as we know it.

Physicists – sharing the Standard Perspective – argue that quantization of GR is necessary, but many of these arguments apply more broadly against any ‘hybrid’ theory of QG in which gravity is fundamentally classical. These arguments include:

- *The universality/fundamentality of quantum theory*: Quantum theory says that every dynamical field must be quantized, and GR tells us that spacetime is a dynamical field, therefore spacetime must be quantized;⁴⁸
- *Unification*: Through quantization, gravity (whether conceived of as the curvature of spacetime or as a force), is treated as quantum in the same way as matter and the (other) fundamental forces; (but note that although quantization is promoted as a way of producing a ‘full’ quantum theory of gravity, rather than a ‘hybrid’ semiclassical theory, it still may not properly represent unification, §4);
- *Spacetime singularities*: The presence of particular singularities in GR could be resolved by quantization of gravity, and points to the need for quantization (this argument is discussed in §6);
- *Heuristic arguments and thought experiments*: Various arguments in which the involvement of quantum matter with classical gravity leads to paradox,

⁴⁷ Butterfield & Isham (2001) also describe how different structures in GR may be quantized.

⁴⁸ This argument is made by many authors, but prominently in Rovelli (2001, 2004).

inconsistency, or otherwise difficult-to-accept conclusions, and are thus taken to indicate the necessity of the quantization of gravity;⁴⁹

While these arguments are widely appealed to, none conclusively establish that QG must be a theory of quantized GR.

A comment on the first of these arguments, regarding the fundamentality of quantum theory: It is a key assumption behind the Standard Perspective on addressing the Primary Motivation for QG, and may be questioned if we take the New Framework Perspective, according to which neither GR nor quantum theory is fundamental as they stand. Additionally, as noted earlier, quantization is also standardly taken as the main characteristic principle of QM to be featured in QG, and thus to act in the role of principle-correspondence, as part of fulfilling the Primary Motivation (this could be true on either the New Framework Perspective, or the Standard Perspective). But, even if we grant the fundamentality of quantum theory, it still does not follow that gravity must be quantized: it may depend on what *interpretation* of QM we are working with. As Adlam (2022) states,

[...] the relevant notion of unification here is the idea that there should ultimately be only one type of stuff – i.e. quantum stuff – and therefore since ordinary matter gravitates, and ordinary matter is assumed to be composed entirely of quantum stuff, it follows that quantum stuff must gravitate. But not all interpretations of quantum mechanics agree that there exists only ‘quantum stuff’, and therefore this conclusion is by no means inevitable. (Adlam, 2022, p. 115)

To recap: while the quantization of gravity is overwhelmingly viewed as the correct way of approaching the problem of QG, it is not the only one. Quantization is taken as a motivation for QG, but is itself driven by assumptions about the fundamentality of quantum theory as well as desires for unification. These principles also underlie the Standard Perspective on addressing the Primary Motivation for QG. Still, neither quantization nor the more-basic principles motivating it are necessary constraints on a theory of QG.

5.1 Hybrid Theories and Emergent Gravity Approaches

If gravity is not quantized, one possibility is that it remains fundamentally classical, as in semiclassical gravity or other hybrid approaches. Recall that LEQG

⁴⁹ Two of these have already been described in §2, from Feynmann, and Bohr & Rosenfeld (1933). Others include Eppley & Hannah (1977); Peres & Terno (2001). These have already been well covered in the literature, and as such I do not go over them again here, Großardt (2021, 2022); Huggett & Callender (2001); Kent (2018); Mattingly (2005, 2006); Oppenheim (2023); Rydving et al. (2021); Tilloy (2018); Wüthrich (2005).

(§3) is an effective field theory for QG (it is taken to reproduce the results of QG at low energies), and comprises both a semiclassical treatment of gravity and a description of the quantum fluctuations of the LEQG field. The dominant attitude towards semiclassical gravity is that it is an approximation to LEQG in the situations where quantum fluctuations of the spacetime geometry can be ignored; that is, semiclassical gravity is a *mean field description* in LEQG. This attitude holds that gravity is fundamentally quantized (and will be described as such by the more fundamental, unknown, theory of QG), and is only effectively treated as classical (i.e., as an approximation in the ‘low-energy’ domains where we know GR is successful).

As Huggett et al. (2023) points out, though, there are two alternative attitudes towards semiclassical gravity. One of these holds that it is not the mean field description of LEQG, but some other theory. If the low-energy approximation of the more fundamental theory of QG were not LEQG, this would allow that QG is not necessarily a theory of quantized gravity. This view fits the various examples of ‘emergent gravity’ approaches mentioned next, as well as proposals such as that of Oppenheim (2023). Huggett et al. (2023) take this view as an ‘epistemically careful’ one, where we accept that semiclassical gravity approximates some theory X , which in turn approximates QG, but we take no stance on whether X is LEQG, or some other alternative (one in which gravity may even be classical – or not). It allows us to remain neutral on the exact relationship between semiclassical gravity and QG. Alternatively, there is the possibility that semiclassical gravity (§2) is the fundamental theory of QG. This attitude is not a popular one, being plagued with both conceptual and mathematical difficulties.⁵⁰

The label ‘emergent gravity’ refers to a broad and eclectic class of approaches to QG which treat gravity as a low-energy effective phenomenon that emerges from the collective behaviour of more fundamental microscopic degrees of freedom. The microphysics itself is not based on a discretization or quantization of GR (Linnemann & Visser, 2018). Simple examples of this approach are analogue models of spacetime from Bose-Einstein condensates (BECs).⁵¹ Others include Sakharov’s induced gravity (Sakharov, 1967, 2000), Jacobson’s gravitational thermodynamics (Jacobson, 1995), Padmanabhan’s gravitational thermodynamics (Padmanabhan, 2010), and Verlinde’s entropic gravity (Verlinde, 2011).

⁵⁰ See fn. 15.

⁵¹ See Barceló et al. (2001); Hu (2009); Liberati et al. (2009); Visser et al. (2002); Volovik (2003). Philosophical discussions: Bain (2008); Crowther (2016).

Typically, these approaches share the conviction that gravity is analogous to hydrodynamics (or thermodynamics), and the Einstein equations are akin to the wave equation for sound in a medium (i.e., they are higher-level equations of state). Accordingly, they hold that quantization of GR does not give us a theory of the fundamental micro degrees of freedom – by analogy, we would arrive at a theory of phonons rather than a description of the underlying ‘atoms’ of space-time. If approaches that quantize the metric tensor produce theories of particles analogous to phonons, then it is unsurprising that they should break down at high-energy. Their breakdown can motivate the search for a high-energy theory beyond (quantum) GR, but we cannot say that the degrees of freedom of the high-energy theory would themselves need to be quantized in order to produce a theory of QG.

5.2 Tabletop Experiments

In spite of the issues discussed earlier, there is widespread agreement that the quantum nature of gravity cannot be established (or ruled out) by conceptual niceties, theoretical arguments or mathematical deduction: it must be confirmed empirically. Even Rosenfeld maintained that the Bohr & Rosenfeld (1933) argument against semiclassical gravity by analogy with ‘semiclassical electromagnetism’ (described in §2) did not demonstrate the inconsistency of semiclassical approaches, because no formal arguments can establish the necessity of quantization. Rather, according to Rosenfeld (1963), only experiment can answer such questions.

And, indeed, since then, various experiments have been proposed and carried out. The Colella-Overhauser-Werner neutron interferometry experiment, touted as ‘observation of gravitationally induced quantum interference’ demonstrated that the gravitational field affects the behaviour of quantum systems (Overhauser & Colella, 1974; Colella et al., 1975). The much-discussed Page and Geilker quantum Cavendish experiment (Page & Geilker, 1981) claimed to provide evidence supporting the hypothesis that a consistent theory of gravity coupled to quantized matter should also have the gravitational field quantized; however, their results were shown to be dependent on the particular interpretation of QM adopted.⁵² These experiments, respectively, show that quantum systems both feel the effect of gravity, and that quantum systems can be sources of gravitational fields – in accordance with predictions of LEQG

⁵² By ‘much-discussed’ here, I mean that there were many confused expressions by physicists as to why Page and Geilker thought it necessary to undertake such an experiment in the first place: the results being thoroughly unsurprising. For philosophical discussion see Huggett et al. (2023); Wallace (2022). Also a nice summary in Kiefer (2007a).

(Adlam, 2022). However, neither of these experiments probe states in which gravity remains in a coherent quantum superposition, and as such, we have no empirical observations in the regime of LEQG where superpositions of spacetimes are important. More recently, a new class of ‘tabletop QG experiments’ have been proposed which claim to do exactly this (Bose et al., 2017; Marletto & Vedral, 2017). Huggett et al. (2023) argue that how we assess the empirical scope of such experiments, though, is ‘a much more intricate affair than may be at first thought, and crucially depends on matters of physical interpretation that are not settled by rational argument’ (p. 74). Note that these experiments are – at this time – proposals, likely years away from being carried out.

6 Singularity Resolution

Widely cited as a motivation for QG is the need to resolve particular space-time singularities in GR. Spacetime singularities are pathologies of a spacetime, and there are various ways in which spacetimes can be singular.⁵³ Next, I discuss the two most common categories of spacetime singularity: incomplete geodesics, and curvature singularities.⁵⁴ Theorems of Penrose and Hawking (Penrose, 1965; Hawking & Ellis, 1973), show that singularities (incomplete geodesics) are unavoidable in GR under very general, physically reasonable conditions. Because a common interpretation of these singularities is as representing the ‘breakdown’ of spacetime, there is the worry, typically expressed, that GR ‘contains the seeds of its own destruction’. Spacetime singularities are thus taken as *internal motivation* for seeking a theory of QG, that is, they come from the theory itself. The corresponding constraint on QG is that it be a theory that *resolves* the spacetime singularities – meaning that it should both be non-singular (in any physically problematic way), and explain ‘what happens’ in those domains where GR is thought to break down. Yet, it is not clear exactly if, how, or why, the spacetime singularities in GR signal a ‘breakdown’ or incompleteness of GR – and thus, whether, or how, they in fact serve as internal motivations, or constraints, for QG.

Although (particular) spacetime singularities are the most commonly cited motivation for QG, they are not the only singularities which could be resolved by QG. In QFT, there are several different types of divergences, including the UV divergences associated with the perturbative nature of the theory (§2.3), and Landau poles. These may be taken as internal reasons for not treating the

⁵³ See, e.g., Curiel (1999); Earman (1996) for a discussion of different types.

⁵⁴ According to Earman (1996), curvature singularities lead to geodesic incompleteness, whereas the opposite is not true. Curiel (1999, §1.1) argues that the two notions are actually independent.

framework of QFT, or particular QFTs (such as the Standard Model), as fundamental, and for seeking QG.⁵⁵ The UV-divergences do not necessarily point to the need for new physics, given that they could just be artefacts of the misapplication of perturbation theory. While the presence of Landau poles in the Standard Model of particle physics presents a stronger internal motivation for new physics, it is one that is usually ignored given that new physics is expected before these scales for *external* reasons, that is, not from within the theory itself.

Here, I follow Crowther & De Haro (2022) in exploring how singularity resolution can motivate QG, and what the implications are for the theory sought. We find that, in general, singularities in current theories do not automatically point to the need for a new theory. While there are at least two examples of singularities (curvature singularities, and Landau poles) that do arguably motivate a new theory, it is not certain that these point to the need for QG rather than a different theory formulated ‘at the level of current theory’ (i.e., a new classical theory of gravity such as a modification of GR, or an axiomatic formulation of QFT that does not feature divergences).

In §6.1–6.2, I introduce the relevant QFT singularities, and §6.3 discuss various responses to these. In §6.4–6.5, I introduce the relevant GR singularities, and §6.6, discuss various responses to these.

6.1 UV Divergences in QFT

In §2.3, I described the problem of divergences associated with the perturbative methods in ‘conventional QFT’ (CQFT), and which lie at the heart of the formalism itself. Historically, these infinities were removed in particular theories, such as quantum electrodynamics (QED), via the procedure of renormalization which rendered the theory finite and (impressively) predictive. This procedure, however, was physically suspicious, and the perturbative approach to QFT itself remained intrinsically approximate and conceptually problematic.⁵⁶

One response was the development of *axiomatic QFT* (AQFT): instead of introducing informal renormalization techniques to treat interactions, this approach attempts to put QFT on a firm, non-perturbative footing, by specifying a mathematically precise set of axioms at the outset. Then, models of the axioms are constructed (constructive QFT). Importantly, this approach is not an attempt at QG, but simply a new formulation of QFT at the level of QFT – that is, as a combination of QM and special relativity without any singularities in

⁵⁵ See, e.g., Ellis et al. (2018); Kiefer (2007b).

⁵⁶ The discussion in this section largely follows Crowther & De Haro (2022), §3.

the theory.⁵⁷ Unlike CQFT, it is not an intrinsically approximate theory, since it is supposed to be directly defined on Minkowski spacetime, and so remain well-defined at arbitrarily small and arbitrarily large length scales. Although there are various simplified toy models satisfying the axioms of AQFT, there have been no realistic models constructed. In particular, it has not been demonstrated that QED or any other successful theories in high-energy physics admit formulations that satisfy the axioms of AQFT.

Meanwhile, in mainstream high-energy physics, the development of RG techniques led to a framework for studying CQFT systems at different energy scales, and ultimately to the discovery of the Standard Model. The RG analysis demonstrates that the means by which the cutoff is implemented has no bearing on the low-energy physics; the only effects that are significant at these scales are changes in the coefficients of finitely many interaction terms (the renormalizable interactions). In other words, RG theory tells us that large-scale phenomenology gives us almost no information about the nature of the short-distance cutoff, if there is one.

The dominant philosophical interpretation of this CQFT picture is that the UV divergences are not a real physical problem, but rather indications of the limitations of the perturbative framework of QFT. The framework itself is taken to be inherently approximate,⁵⁸ and its models are supposed to be effective theories: not to be valid to arbitrarily high energy scales. ‘This, in essence, is how modern particle physics deals with the renormalization problem:⁵⁹ it is taken to presage an ultimate failure of quantum field theory at some short length scale, and once the bare existence of that failure is appreciated, the whole of renormalization becomes unproblematic, and indeed predictively powerful in its own right’ (Wallace, 2011, p. 120). The idea is that, whatever the unknown physics of QG turns out to be, the success of the CQFT models at known energies is explained, thanks to the RG.

This interpretation of CQFT as *effective* means that the theory is not supposed to be reliable at short length scales. In particular, the need to employ a short-distance cutoff is not taken to indicate *anything* regarding the physics beyond. This is in contrast with the case in condensed matter physics, where the RG is also employed because the system is described by a theory which diverges in the UV, but in which case the divergences, and the need to employ a short length scale cutoff is consistent with the existence of something physical – we know that we cannot treat matter as continuous at arbitrarily short length

⁵⁷ See, e.g., Fraser (2011).

⁵⁸ Cf. Fraser (2020).

⁵⁹ Footnote in the original suppressed.

scales, because matter has a discrete structure at the atomic scale. In CQFT, however, there is no empirical evidence for the existence of a real physical cutoff (e.g., a discrete structure for spacetime at extremely high energies). It is possible that the UV divergences in CQFT simply reflect limitations of the theory, rather than any new physics.

6.2 Landau Poles

A motivation for treating particular QFTs (rather than the framework of QFT itself) as effective field theories is the existence of Landau poles. This type of infinity is thought more concerning than the UV-divergences because it is not taken to be merely due to the limitations of perturbative analysis.

The most familiar example of Landau pole is in QED. This theory is renormalizable, so in principle it should be able to be extended to arbitrarily high energies. Yet, the renormalized coupling grows with energy scale, and becomes infinite at a finite (though extremely large) energy scale, estimated as 10^{286}eV (the original result comes from Landau et al. (1954)). The existence of this ‘pole’ could mean the theory is mathematically inconsistent. This is avoided if the renormalized charge is set to zero, that is, if the theory has no interactions. There is indication that in QED, the renormalized charge goes to zero as the cutoff is taken to infinity (a physical interpretation of this is that the charge is completely screened by vacuum polarization). This is a case of quantum *triviality*, where quantum corrections completely suppress the interactions in the absence of a cutoff. Since the theory is supposed to represent physical interactions, the coupling constant should be non-zero, and so the Landau pole and the associated triviality might be interpreted as a symptom of the theory being effective, or incomplete (i.e., that it fails to take into account other fundamental interactions relevant at high energy scales).

QED and ϕ^4 theory are thought to be trivial in the continuum limit in this way. In other words, RG analysis of QED and ϕ^4 does not indicate that these theories possess a stable UV fixed point (Lüscher & Weisz, 1987; Gies & Jaeckel, 2004). This means that the Standard Model of QFT is thought to suffer Landau poles both for the electron charge, and the Higgs.⁶⁰ Although the Landau pole in QED is problematic for the theory, it is usually ignored because it concerns an energy scale where QED is not thought to be valid anyway, given that electroweak unification occurs at an energy scale lower than this. The Higgs triviality is

⁶⁰ However, since the Landau pole in QED is normally identified through perturbative one-loop or two-loop calculations, it is possible that the pole is a sign that the perturbative approximation breaks down at strong coupling. Similarly, although the Higgs also is suspected to be trivial at high energy, this has been difficult to prove theoretically.

thus more interesting for the Standard Model, as potentially indicative of new physics at high energy.⁶¹ However, it also concerns an energy scale where *QFT itself* is not thought to be valid, based on the *external* arguments for the Planck length.

In summary, the Landau pole divergences in CQFT are typically interpreted as part of the formal (mathematical) grounds – internal to the theory – for treating particular QFTs as effective. But there are also external grounds (not from the theory itself) for treating particular QFTs – and perhaps the framework of QFT itself – as effective, which come from the motivations for QG. These external motivations hold regardless of whether or not there are divergences inherent to QFT, but they are reasons why one might not be concerned with finding a singularity-free theory of QFT in order to describe the world at arbitrarily small length scales.

6.3 Responses to QFT Singularities

There are three different types of divergences associated with QFT that may bear on the problem of QG: non-renormalizability of gravity (§2.3), UV-divergences in QFT, and Landau poles. The perturbative non-renormalizability of quantum GR (LEQG, §2.3) does indicate the need for a theory of QG valid at the Planck scale, though this theory may be non-perturbative quantum GR (i.e., the asymptotic safety scenario). The UV divergences of QFT may be artefacts of perturbation theory, and there is a danger of misinterpreting the need for a high energy cutoff as indicating a physical or operational minimal length. They do not point directly to the need for QG. Landau poles in the Standard Model of particle physics do (likely) represent an incompleteness of the theory and motivate the need for a new theory – though this new theory need not be QG. Landau poles, however, are typically ignored by physicists who already expect QG to replace QFT at the Planck scale for external reasons (§3.1, also Held (2019, §1)).

From the preceding discussion, we can identify four different possibilities in response to these singularities (following Crowther & De Haro (2022)):

- i. *AQFT view*: Singularities in CQFT motivate a different QFT framework, one whose theories are singularity-free, but which does not include gravity (as in AQFT);
- ii. *New physics*: Singularities in CQFT motivate a new, more fundamental theory at high-energy, and motivate treating our current theories as effective

⁶¹ This issue is also related to the principle of *naturalness*, which the Higgs mass conflicts with.

(applicable only at low energy scales), consistent with external motivations for QG;

- iii. *Effective theory view*: Ignore the Landau poles in the Standard Model and perturbative non-renormalizability of gravity (we shouldn't worry about interpreting them), since we have external reasons for thinking of these as non-fundamental effective theories, to be replaced by QG at high-energy scales (i.e., we appeal only to the external motivations for new physics, and the singularities in current theories do not count as motivations for new physics);
- iv. *Asymptotic Safety view*: UV divergences due to perturbative non-renormalizability and Landau poles do not motivate new physics or a new theory according to the asymptotic safety scenario for gravity and the Standard Model; these singularities do not appear in the full (non-perturbative) theory.

There is one more response to singularities in CQFT, which has been expressed prominently by Jackiw (1999, 2000) and Batterman (2011).

- v. *Emergent physics view*: Singularities in CQFT are of physical significance, but not motivation for new, more fundamental physics. The singularities are important for facilitating and understanding the emergent, low-energy physics.

Following Crowther & De Haro (2022), we can classify these positions under four different 'attitudes' towards singularities. *Attitude 1* means that the singularities are resolved 'at the level of QFT', without pointing to a more-fundamental theory or framework. Here, (i) and (iv) fall under Attitude 1. *Attitude 2* means that the singularities are to be resolved by a more fundamental theory of QG, and here (ii) is an example. *Attitude 3* is a positive attitude towards singularities, which means not resolving them at any level, because we have reason to keep them (or reasons why they cannot be removed). Example (v) is a case of Attitude 3, where the singularities are seen as importantly explanatory, or 'essential'. Finally, *Attitude 4* means that we do not care about the singularities in question; they are of no significance. (iii) is an example of this attitude. Importantly, one can (and should) take a different attitude towards different types of singularities. This classification also holds when considering spacetime singularities: taking them as pointing to QG is only one possible option (depending on which singularities are being considered).

The QFT divergences, in contrast to the spacetime singularities to be discussed, are not typically appealed to as motivation for QG (though this argument occasionally appears); more commonly the resolution of these is taken as

a means of confirmation, or indication of pursuit-worthiness of approaches to QG that resolve these divergences (though we still classify this falling under Attitude 2). The adherence to quantization as the means of solving the problem of QG is consilient with the resolution of QFT divergences: if we quantize gravity, the UV divergences are usually expected to be resolved ‘for free’ (via a minimal length), and this is taken to support approaches to QG that quantize GR. As stated in §2.3, the (alleged) resolution of the UV-divergence of quantum GR by string theory is presented as one of the selling points of string theory.

Some believe that the singularities in GR and QFT are to be cured by the existence of a minimal length, even without quantization of gravity (note, too, that the minimal length need not represent an actual discretization of spacetime, but may be an operational minimal length, e.g., due to an extended probe).⁶² Henson (2009) expresses this view, particularly in regard to the need for QG to resolve the divergences associated with the non-renormalizability of GR, and uses it as motivation for the discreteness postulated by an approach to QG known as causal set theory.

6.4 Geodesic Incompleteness

The definition of spacetime singularities in terms of *geodesic incompleteness* states that a spacetime is singular if and only if it contains an incomplete, inextendible timelike geodesic. Such a geodesic is the worldline of a freely falling test object; the property that makes it singular is that the worldline ends within finite proper time and cannot be further extended. While this definition forms the basis of the Penrose and Hawking singularity theorems, it is not without problems (see Curiel, 2023, §1.1). Geodesic incompleteness appears to be a genuine physical worry, because it means that ‘particles could pop in and out of existence right in the middle of a singular spacetime, and spacetime itself could simply come to an end, though no fundamental physical mechanism or process is known that could produce such effects’ (Curiel, 1999, p. S140). Geodesic incompleteness leads to a lack of predictability and determinism, and so could indicate that the theory is incomplete (Earman, 1995, §2.6).⁶³ If the breakdown of determinism were visible to external observers, ‘then those observers would be sprayed by unpredictable influences emerging from the singularities’

⁶² For example, Ellis et al. (2018) states there is a widespread sentiment among QG physicists, that the singularities in GR and QFT are due to the assumption of a spacetime continuum. For more on the minimal length, see Hossenfelder (2013).

⁶³ Basically, determinism is said to hold if specifying the state of the world at a time together with the laws of nature jointly determine the state of the world at all times. In GR, physicists typically take determinism to hold if the spacetime admits of a well-posed initial value formulation. Cf., Doboszewski (2019); Smeenk & Wüthrich (2021) for recent work on the problems of defining determinism in GR.

(Earman, 1992, p. 171). This would represent a nasty form of inconsistency – as Earman puts it, the laws would ‘perversely undermine themselves’.

The breakdown of determinism inside black holes occurs beyond the *Cauchy horizon* (inside the event horizon): beyond this surface, the Einstein equations no longer give a unique solution. In response, Penrose (1979) proposed *strong cosmic censorship* (SCC), which postulates that the appearance of the Cauchy horizon in Schwarzschild black holes is non-generic, and that the interior region of these black holes is in some way unstable (under small perturbation of initial data) in the vicinity of the Cauchy horizon. Any passing gravitational waves would prevent the formation of Cauchy horizons, meaning that instead, space-time would terminate at a ‘spacelike singularity’, across which the metric is inextendable.

SCC ensures that no violations of predictability are detectable even by local observers (i.e., an astronaut on a geodesically incomplete worldline would detect nothing up until, and presumably after, her disappearance), and so, the truth of this conjecture would render any singularities (incomplete geodesics) harmless in regard to determinism. As Dafermos & Luk (2017, p. 5) states, ‘The singular behaviour of Schwarzschild, though fatal for reckless observers entering the black hole, can be thought of as epistemologically preferable for general relativity as a theory, since this ensures that the future, however bleak, is indeed determined’. Thus, SCC may be able to save GR from the charge of incompleteness.

There are two issues that have been raised in the philosophy literature that bear on the question of indeterminism and incompleteness in GR, however. The first regards the nature of the spacetime singularities themselves: (essential) singularities are not located at a *point* of spacetime (if a singularity is located at a point in spacetime, then this is indicative of having a non-essential singularity, that is, one that can be removed, in the case of incomplete geodesics, by extending the geodesics beyond that point). As Curiel (1999, 2023) emphasizes, a singular spacetime does not have any ‘missing points’ of spacetime. (For this reason, Curiel and others have stressed the *global*, rather than local nature of singularities.) Earman (1995, 1996) also uses this argument against the claim that spacetime singularities indicate that GR (alternatively abbreviated as GTR) is incomplete,

In GTR the metric field g_{ab} is now a physical field whose singularity behavior we have to judge. And the judgment has to start from the fact that a general relativistic spacetime M, g_{ab} is such that g_{ab} is defined and differentiable at every point of M —there are no singular points of spacetime where the laws of GTR fail to apply.

This definitional move may seem facile. In the case of a closed Robertson-Walker spacetime, Brandenberger *et al.* see a kind of incompleteness resulting from the initial and final singularities. '[T]he [final] singularity implies that we cannot answer the question what will happen after the "big crunch" or (in the case of an expanding universe) what was before the "big bang"'.⁶⁴ In response I would pose a dilemma. Either the initial and final singularities are essential—in the sense that it is impossible to extend through them in a way that EFE make sense [...] – or not. If so, then by the lights of GTR, talk about 'before' the big bang and 'after' the big crunch is physically meaningless. If not, then GTR can say something about the before and after. In either case GTR does not stand convicted out of its own mouth of raising meaningful questions it cannot answer. (Earman, 1996, pp. 631–632)

Smeenk (2013) makes the same argument in regard to the big bang singularity: that the laws of GR apply throughout the entire spacetime, and there is no obvious incompleteness. If GR is the correct final theory, 'then there is nothing more to be said in regard to singularities' (p. 634). While Smeenk does not view spacetime singularities as motivation for seeking QG, he states that there are other reasons for doubting that GR is the correct final theory, and good reasons to expect that the successor to GR will have novel implications for singularities – that is, he recognizes *external motivations* for QG, without counting the GR singularities as internal motivation for seeking a new theory.

Although Earman generally advocates a 'tolerance for spacetime singularities' – arguing that we can treat them as *predictions*, rather than pathologies of GR – he nevertheless believes there is one way in which the charge of incompleteness may be justified. This is the idea, described earlier, that if SCC does not hold, then the determinism of GR is undermined (Earman, 1995, 1996). Earman ties the determinism of GR to spacetime models that are globally hyperbolic and thus admit of a globally well-posed initial value formulation. The problem with this, however, is a 'dirty open secret', that determinism in GR fails without help by fiat (i.e., the imposition of ad hoc constraints that simply rule out those spacetimes that do not admit a globally well-posed initial value formulation).

This leads to the second issue discussed in the philosophy literature – whether GR is a deterministic theory, and whether all models of the theory represent physically possible spacetimes. These questions are explored in Smeenk & Wüthrich (2021), which also highlights a tension between the 'philosopher's conception' of determinism and the conditions needed for having a well-posed initial value formulation in GR physics. Doboszewski (2019, 2020) also discusses the problems of defining determinism in GR, with the

⁶⁴ (Brandenberger et al., 1993, p. 1629).

former paper arguing for a pluralistic conception. These issues with determinism in GR may take the bite out of the worry that geodesic incompleteness without SCC is a problem motivating QG (though, of course, this depends on one's convictions regarding the need for a deterministic theory).

6.5 Curvature Singularities

According to the definition of curvature singularities, a spacetime is singular if the curvature, especially the scalar quantities constructed by contracting powers of the Riemann tensor, grows without bound in some region of the spacetime.⁶⁵ This gives rise to various problems, such as unbounded tidal forces and the lack of consistency of the semiclassical approximation. The semiclassical approximation being referred to here treats GR as an effective field theory, with the Einstein-Hilbert action supplemented with additional higher-curvature terms that represent the quantum corrections to the theory.⁶⁶ If we take a model, M , of GR with a curvature singularity, and then check whether it is an approximate solution of the equations with the quantum corrections, we find that in general, M may be a good solution far away from the singularity, but as we approach the singularity, the higher-order (high curvature) terms will start to dominate over the lowest order term (Einstein tensor). Because of this, M will not be a model of the quantum-corrected theory. This shows how the curvature singularities can be used to predict that GR 'breaks down', since it reveals an inconsistency between GR and (expected) QG effects that manifest at high curvature in the region close to the curvature singularity.

This is a standard argument found in the physics literature, and physicists tend to find the curvature singularities more concerning than geodesic incompleteness, which has been the focus in the philosophy literature. Notice, however, that here it is not the singularity *itself* that is the problem (depending on the outcome of the charge of indeterminism described earlier), but the *expected quantum effects* in regimes of high curvature – it is this that motivates the need for a new theory.

6.6 Responses to Spacetime Singularities

Earlier, I argued that geodesic incompleteness may not represent an incompleteness of GR, depending on the status of the SCC and/or the assessment of (in)determinism of GR. Curvature singularities, meanwhile, are arguably problematic insofar as they represent domains of extreme curvature where GR is no

⁶⁵ There are also curvature singularities whereby some of the physical components of the Riemann tensor do not have a limit; see Earman (1995, p. 37).

⁶⁶ See Crowther & De Haro (2022), §2.3 for full details.

longer thought to be reliable, and thus do motivate the need for a new theory to replace GR. If we take it that GR is unreliable due to its neglect of quantum effects in these domains (as is the standard interpretation), then this new theory required is QG.

The dominant attitude towards spacetime singularities – especially the big bang singularity and black hole singularities – is as motivation for QG. This attitude treats GR as a classical approximation, an ‘effective theory’, and the singularities in GR signal its limitations as such. Accordingly, singularities should be resolved by a more fundamental theory of QG.⁶⁷ This attitude is inspired by the history of physics, in particular the suggestion that singularities in classical theories signal the breakdown of these theories and the need for quantization. An oft-cited example in this context is the resolution of the instability of the hydrogen atom in classical electrodynamics, by quantum mechanics. Electrodynamics predicts that the electron radiates energy as it orbits the nucleus and eventually falls into the nucleus, at which point the force becomes infinite. QM solves this by, first, rejecting the classical picture of an electron ‘orbiting’ around the nucleus. Second, by confining the electron to discrete energy levels, and allowing it to emit energy only in discrete packets of energy (photons). And finally, the lowest energy level allowed by the theory is where the electron is, on average, located at a finite distance from the nucleus (the ‘Bohr radius’), so that it can never fall in. In regard to the spacetime singularities, the idea is that, analogously, gravity should be quantized and that QG solutions have a discrete spectrum, similar to the discrete spectrum of QM – and that, accordingly, singularities can be avoided.

While this argument by analogy appears often in the literature, it may be criticized on various grounds. Earman (1996) argues against the applicability of this analogy on two points (1) the instability of the atom in this picture comes not from a single, well-confirmed theory, but by combining the Rutherford model of the atom with classical electrodynamics, (2) this combination leads to a prediction that is empirically ‘falsified by the fact that the material world exists’ (p. 629). By contrast, Earman states that there is no such definitive conflict with observation in the case of GR. The predictions of GR regarding the formation of black holes in gravitational collapse are meeting with increasing empirical success (though there is more difficulty in assessing the empirical success of Standard Model of cosmology which gives us the big bang singularity). Another line of criticism is expressed by Tilloy (2018): just because

⁶⁷ What exactly is meant by ‘resolution’ in different contexts is an interesting philosophical question. Thébault (2023) explores this question in the context of big bang singularity resolution in quantum cosmology.

quantization might solve the problem of singularities in GR does not mean it is the only solution, and promoting it as such is, in a way, avoiding the problem by saying it will be dealt with by an unknown, more-fundamental theory. Indeed, by looking to QG rather than seeking to understand singularities in GR ‘at the level of GR’, we may be missing key physical insights of the theory. This attitude is advocated by Earman (1996) (following Misner (1969)), who argues that singularities in GR should be treated as *predictions* of the theory.

In Crowther & De Haro (2022), we find there are four different attitudes one may take towards singularities in GR, the same as those mentioned earlier for the divergences in QFT. *Attitude 1* for spacetime singularities means that they are to be resolved classically, or ‘at the level of GR’, rather than pointing to QG. *Attitude 2* means they are to be resolved by QG; this is the attitude that dominates the physics literature, and which I have focused mostly on here. *Attitude 3* is a tolerance for spacetime singularities, which means not resolving them at any level, because we have reason to keep them. This is exemplified by Earman’s attitude expressed earlier, where they are treated as predictions of GR, not pointing to QG. *Attitude 4* is an indifference to spacetime singularities: they are not of any significance. We find examples of each of these attitudes in the literature.

The choice of which attitude to take depends on the singularity being considered. We can – and should – adopt different attitudes towards different singularities. The choice also depends, amongst other things, on one’s disposition towards the internal versus external motivations for QG, and one’s position in regard to scientific realism. For instance, if the spacetime singularities (geodesic incompleteness) are not necessarily problematic for GR, then these would not count as internal motivations for QG. By contrast, in QFT the Landau poles are more readily interpreted as signalling incompleteness of the theories; these are, however, typically dismissed because they occur in regimes where the theory is not thought to be applicable anyway *for external reasons* (e.g., the existence of the Planck scale as a minimal length). The tension between the internal and external reasons for treating a theory as effective is reflected in differing possible attitudes towards singularities in current theories: whether they are to be resolved by a more fundamental theory of QG, whether we should instead look to resolve them ‘at the level of current theories’ (i.e., through developing a new framework for QFT, or modifying GR, etc.). Those who put weight on the external motivations for QG may tend to disregard alternative, internal, possibilities for singularity resolution.

Even if singularity resolution is not, in all cases, a motivation necessitating QG, it can be used as a *guiding principle* inspiring new work. And, if a new theory of QG naturally resolves the problematic spacetime singularities (i.e., if

QG describes the ‘domains’ where GR is thought to ‘break down’, or be problematically in deterministic, as well as explains why these occur in GR, and the theory itself not feature any problematic singularities) then, just as in the case of the QFT divergences, this may be promoted as a *means of confirmation*, or an indication of pursuit-worthiness. Making the case for a stronger role, where singularity resolution is taken as a *criterion of acceptance* for any theory of QG, however, requires further argument.

7 Black Hole Thermodynamics

The area of research known as *black hole thermodynamics* (BHT) is currently one of the most important in theoretical physics; it suggests, tantalisingly, at a deep connection between GR, QFT and thermodynamics (the ‘three pillars perspective’, §3.4). BHT lacks empirical access or support, but is the most widely accepted, deeply trusted set of purely theoretical propositions in physics.⁶⁸ BHT stems from attempts to bring together GR and QFT, but motivates the Primary Motivation, in suggesting the need for a new theory for its explanation. In doing so, it leads to various desiderata and potential constraints on the new theory, which I discuss in this section.

BHT is formulated in the framework of semiclassical gravity, whose basis is the semiclassical Einstein equations, which couple classical relativistic space-time to quantum matter fields (§2, eq. 4). As mentioned earlier, the dominant attitude towards this framework is that it is an approximation to LEQG in the situations where quantum fluctuations of the spacetime geometry can be ignored. Yet, semiclassical gravity, while thought to provide the most secure theoretical clues towards QG through BHT, is itself bedevilled with a plethora of theoretical and conceptual difficulties (*even when* considered as an approximation, rather than a fundamental theory of QG). Because of this, and the absence of empirical support, the amount of trust in BHT as offering insights into the nature of QG may seem a striking – perhaps even suspicious – situation.

Here, I first introduce the basics of BHT and black hole statistical mechanics; §7.1 discusses Hawking radiation, the theoretical discovery of which was key to accepting BHT; §7.2 discusses the black hole information loss paradox, which remains a hotly debated topic in physics though its relevance for QG is unclear; and §7.3 entropy bounds and the holographic principle, which, arguably, offer more secure avenues along which to seek insights into QG.

Classically, there exists a formal analogy between the laws of black hole mechanics and the familiar laws of thermodynamics. Stated simply, the laws of *black hole mechanics* (BHM) are as follows,

⁶⁸ Curiel, fn. 15.

- 0 **Zeroth Law of BHM:** For a stationary black hole, the surface gravity κ of the event horizon is constant everywhere on the horizon.
- 1 **First Law of BHM:** Changes in the mass, M , of a stationary black hole are related to changes in the horizon area, A , angular momentum, J , and electric charge Q as

$$dM = \frac{1}{8\pi} \kappa dA + \Omega dJ + \Phi dQ \quad (11)$$

Second Law of BHM (aka the Area Theorem): The area of the event horizon, A , will not decrease in any physical process:

$$dA \geq 0 \quad (12)$$

where a physical process is any process which satisfies the null energy condition ($T_{\nu\mu} n^\mu n^\nu > 0$ for null n^μ) and during which the event horizon already existing does not ‘vanish’.

- 3 **Third Law of BHM:** No finite series of operations may reduce the surface gravity of a black hole to zero.

The formal analogy holds when the surface gravity is taken to stand in as thermodynamic temperature, and the horizon area as thermodynamic entropy. However, the formal analogy is now generally believed to be not an analogy at all: rather, it is thought that black holes *are* thermodynamic systems.⁶⁹ This change in attitude came following the (theoretical) discovery of the *Generalised Second Law*, and of Hawking radiation (§7.1), which gave reason to attribute non-zero entropy and temperature to black holes.

Classically, black holes have temperature of absolute zero, and do not possess entropy; according to the *No-Hair Theorem*, they are completely characterized by only three parameters: mass, angular momentum, and electric charge. This leads to the worry that the second law of thermodynamics could potentially be violated in black hole spacetimes. For example, if a package of entropy, such as a cup of coffee, is lost into a black hole, and there were no corresponding increase in entropy of the black hole to compensate for this loss, then the total entropy of the universe (or, more correctly, the total entropy of the black hole plus the relevant external systems) has decreased, and the law has been violated. (There is a second way in which we can imagine the second law being violated, via the *Geroch process* (Bekenstein, 1972): see §7.3.) This led Bekenstein (1974, 1973, 1972) to state that the second law, in the presence of black holes, is ‘transcended’.

⁶⁹ For a much more systematic and thorough account, see Curiel (2023); Wallace (2018a,b). Other useful recent reviews: Almheiri et al. (2021); Harlow (2016).

In response, and by reference to the *Area Theorem* (Hawking, 1971), Bekenstein suggested that black holes possess entropy, S_{BH} , and that this is proportional to A . Following the (theoretical) discovery of Hawking radiation, which states that black holes radiate with temperature $\kappa/2\pi$ (Hawking, 1975), the constant of proportionality has been set, so that,

$$S_{BH} = \frac{kAc^3}{4G\hbar} \quad (13)$$

(where k is Boltzmann's constant). This is known as the Bekenstein-Hawking entropy formula. Bekenstein also proposed that the (original) second law of thermodynamics be replaced by the *Generalised Second Law* (GSL), which states that the total entropy, S_{tot} , including S_{BH} as well as matter entropy, S_{matt} cannot decrease,

$$dS_{tot} = (dS_{BH} + dS_{matt}) \geq 0 \quad (14)$$

A 'dizzying variety of proofs' have since been produced for the GSL in different regimes, relying on very different assumptions about the nature of the systems described, as well as different mathematical techniques (Wall, 2009). As Curiel (2023, §5.4) points out, though, the contradictory assumptions being made across the different proofs may actually threaten the result as tautology, rather than strengthening it as an interesting principle of nature through the consilience of derivations (the derivations, he argues, are *not* consilient).

Nevertheless, the GSL and Hawking radiation results complete the description of a black hole as an ordinary thermodynamic system, rather than merely being analogous to one. Classically, the analogy only worked so long as heat exchanges with other systems were neglected: according to GR, nothing can escape a black hole, it cannot radiate and cannot be assigned a non-zero temperature. Hawking's result relies on *quantum effects* to show that black holes radiate and thus can be assigned non-zero temperature and entropy: it provides a method by which black holes can be put in thermal contact with each other and with other thermodynamic systems.

Given that black holes can be described as thermodynamic systems, and that all other thermodynamic systems are known to possess an underlying statistical mechanical description, we may speculate that the thermodynamics of black holes, too, is underpinned by a statistical mechanical description. This would imply that a black hole of given mass, charge, and angular momentum, has $\approx \exp(A/4G)$ microstates. The statistical mechanical entropy has been calculated in the framework of LEQG, and reproduces the entropy formula, providing evidence that the black hole thermodynamic entropy has an

associated statistical mechanical description.⁷⁰ Since then, the recovery of the entropy formula by calculations of statistical mechanical entropy (counting of microstates) has been taken as a *guiding principle*, a *means of confirmation*, and – most conspicuously – a prominent *criterion of theory acceptance* for QG. Various approaches to QG have managed to reproduce this result, for different types of black holes, and to varying levels of precision, including in loop quantum gravity (Rovelli, 1996; Meissner, 2004), and in the context of the AdS/CFT duality (Witten, 1998; Aharony et al., 2004). However, the most-discussed results are in string theory: Strominger & Vafa (1996) showed that the statistical mechanical entropy of a certain type of black hole (known as ‘extremal’ black holes) can be calculated in string theory, and, to leading order, the result matches the area formula. This calculation quickly gained widespread acceptance and is viewed as one of string theory’s main successes.⁷¹

The derivation of the Bekenstein-Hawking entropy formula by the counting of microstates is now held as one of the main criteria of theory acceptance for any theory of QG. This is particularly striking given that it is a purely theoretical result that lies far from empirical scrutiny. (Compare with the recovery of GR, the other main criterion of acceptance for QG, which is explicitly motivated by empirical concerns.) This raises interesting philosophical questions regarding the scientific method, and the future trajectory of fundamental physics (Curiel, 2023, §5.3).

The results of BHT, however, are certainly tantalizing, and, if we take them seriously, do seem to point to the need for a new, more fundamental, theory. This is because there is not a way in GR, nor in semiclassical gravity, to describe a black hole as system whose physical properties arise as gross statistical measures over underlying microstates. Additionally, while Hawking radiation is predicted in the framework of semiclassical gravity (more strictly, *QFT in curved spacetime*, see further), this radiation is not generated by micro-degrees of freedom (the event horizon is treated as a simple, classical geometrical structure). A more fundamental theory appears to be required if we are to explain (the origin of) BHT. Because BHT relies on both GR and QFT effects, its explanation serves as an *external* motivation for QG: it is, in essence a problem ‘of our own making’, rather than one suggested by our best theories as they stand on their own.

Nevertheless, Wallace (2018b, p. 116) argues that the convergence of these various different calculations is ‘overwhelmingly’ evidence in support of BHT as underpinned by statistical mechanics – and, furthermore, that this supports

⁷⁰ See Wallace (2018b) for details and original references.

⁷¹ For details, see De Haro et al. (2020); van Dongen et al. (2020); Wallace (2018b).

both the belief in the correctness of LEQG (‘that general relativity can be analysed by QFT at low energies just like any other field theory’) and the correctness of the AdS/CFT duality (‘so that black hole statistical mechanics is dual to the statistical mechanics of a conformal field theory in at least some important cases’). Wallace’s attitude towards the consilience of these derivations thus stands in stark contrast with Curiel’s worries in the case of the (‘multiplicity and multifariousness’ of the derivations for both the) GSL and Hawking radiation. One reason for this may relate to the inter-theory relations of correspondence and reduction that are expected to hold between QG and LEQG: since QG is expected to be a UV completion of LEQG, it is expected that the two theories share approximately the same results in the domains where they overlap, and because QG is also supposed to be more fundamental (analogous to statistical mechanics as more fundamental than thermodynamics), it is also expected to *explain* the BHT results in LEQG. Of course, as mentioned earlier, this idea of correspondence cannot – at this stage, however – be an application of the Generalised Correspondence Principle, given that neither of these theories is among our best accepted (empirically successful) physics. Thus, as suggested by Huggett et al. (2023), earlier (§5.2), we might still choose the ‘epistemically careful’ attitude towards LEQG, as well as the specific relationship between QG and semiclassical gravity.

7.1 Hawking Radiation

Hawking (1974, 1975) showed that the QFT vacuum state becomes a state with (real) particles present if there is a black hole – that is, that black holes *radiate*. More precisely, Hawking considered the vacuum state of a scalar field in the presence of an event horizon and found that it evolves into a thermal state of temperature T . Because the thermal spectrum of the radiation encodes a Planckian temperature that is proportional to the surface gravity, and the surface gravity plays the role of temperature in the zeroth and first laws of BHT, the temperature is associated with the black hole. Whereas, classically, a black hole is thought to absorb everything and emit nothing (including light), Hawking’s result shows that black holes emit thermal radiation with the Planckian power spectrum characteristic of a perfect blackbody at a fixed temperature. It results in energy being radiated away, leading to the eventual ‘evaporation’ of the black hole. A primordial black hole of mass 10^{12} kg, produced in the early universe, would evaporate around now. The Hawking temperature of a solar mass black hole is 10^{-7} K and its lifetime is 10^{64} years. Figures 5 and 6, from Almheiri et al. (2021), depict this process.

A non-technical heuristic explanation attributes the Hawking effect to vacuum fluctuations leading to the separation of positive and negative modes

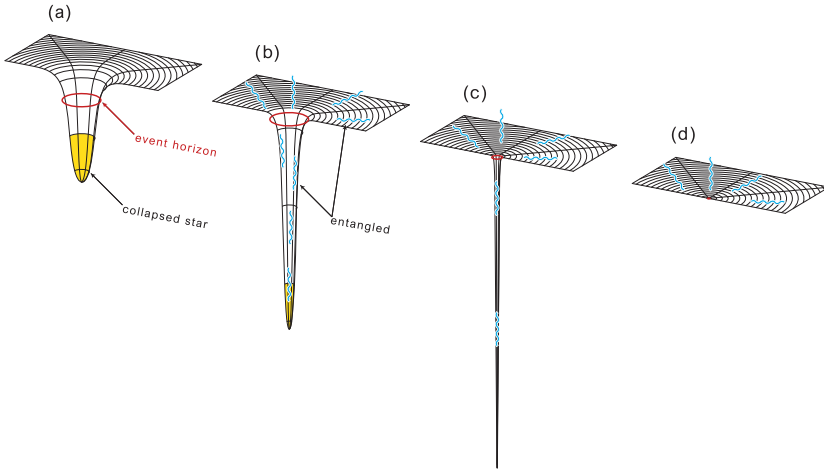


Figure 5 *The Hawking process of black hole evaporation:* (a) Stellar collapse: the interior geometry continues to elongate in one direction, while pinching towards zero size in the angular directions; (b) Hawking process creates entangled pairs of particles, with one partner trapped behind the horizon and the other escaping to infinity where it is observed as blackbody radiation. The black hole shrinks as its mass is slowly carried away; (c) Eventually, angular directions shrink to zero, as does the event horizon; (d) Final stage: a smooth spacetime containing thermal radiation but no black hole.

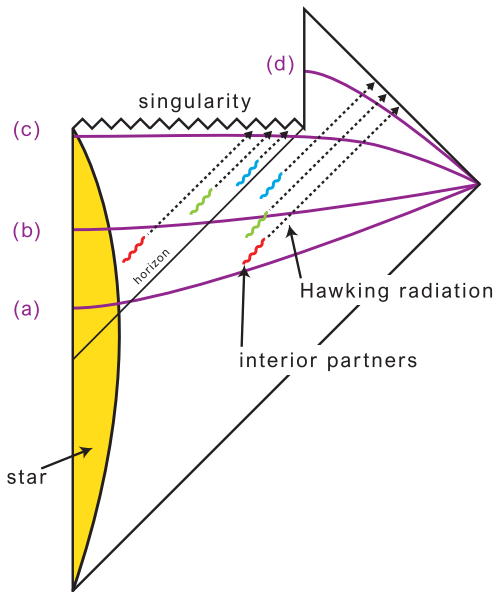


Figure 6 Penrose diagram for the formation and evaporation of a black hole. Spatial slices (a)–(d) correspond to the slices drawn in Figure 5. (Almheiri et al., 2021).

through the black hole's event horizon. The QFT vacuum state swarms with pairs of particles that are being constantly created and annihilated. If such a pair forms near the event horizon, one of the entangled partners may fall into the black hole, while the other escapes to infinity – the separation of the pair precludes their otherwise typical recombination, and results in the exterior region having an effective excess energy (real particles). The escape of the excess particles to the asymptotic region is then what gives rise to outgoing radiation.⁷²

Hawking's original calculation was done using the framework of *QFT in curved spacetime*.⁷³ This is a non-trivial, but consistent, theory of quantum fields on the background of a classical relativistic spacetime, which differs from semiclassical gravity in its neglect of the backreaction of matter on the curved spacetime. This neglect is justified in general if the relevant length scales are large and curvature effects small. Thus, we expect this framework to naturally apply in the cases where the curvature of spacetime is well above the Planck length, and where we can assume (i.e., we have some theoretical grounds for expecting) that we can ignore any quantum properties of spacetime itself. Nevertheless, the empirical support for QFT in curved spacetime is both very indirect and highly nonspecific, since it comes not from indirect measurements concerning the predictions of particular QFTs in curved spacetime, but rather from claims that particular QFTs in curved spacetime show the right behaviour in *various limits*.⁷⁴ An example is the geometric optics limit of electrodynamics in curved spacetime, which is a limit of *quantum* electrodynamics in curved spacetime, describes light rays as tracing out null geodesics given sufficiently high frequency relative to the curvature scale. That light indeed moves on null geodesics in curved spacetime has been observationally verified, for example, in gravitational lensing effects (see Dyson et al. (1920)). The prediction of Hawking radiation is one of only two main specific predictions that could serve as a test for this framework.⁷⁵

Another problem with Hawking's original derivation is that it involves field modes whose wavelengths become shorter than the Planck length near the black hole horizon – this 'trans-Planckian' regime is beyond the domain of applicability of QFT in curved spacetime, and so we should not expect to trust its results

⁷² Note, however, that while this is the standard heuristic picture described, it is quite imprecise, and does not, for instance, allow for a distinction between the Hawking effect and the Unruh effect, cf. Barbado et al. (2016).

⁷³ See Wald (1994).

⁷⁴ This means that same-limit-behaviour might be reproducible by myriad theories other than QFTs in curved spacetime.

⁷⁵ The other is the prediction of a specific primordial density perturbation spectrum associated with cosmic inflation scenarios. For criticisms of this as a possible confirmation of QFT in curved spacetime, see Crowther et al. (2021).

there (instead, requiring a theory of QG to describe this regime). The problem arises due to an exponential gravitational redshift assumed to occur at the horizon, which means that the black hole radiation detected at late times (i.e., the outgoing particles) corresponds to extremely high (divergent) frequency when traced back to the horizon.

Yet, there are at least four other derivations of the Hawking effect in addition to Hawking's original method (Wallace, 2018a). Together, Wallace claims, these strongly suggest that Hawking radiation is a real consequence of QFT in curved spacetime. Curiel (2023), on the other hand, argues that not all these derivations are consilient,⁷⁶ given that some appear to rely on assumptions that contradict assumptions made by others: leaving the physical content of the prediction of Hawking radiation obscure at the foundational level. Curiel argues against identifying Hawking radiation with blackbody radiation given that, in the semiclassical picture, it is not generated by micro-degrees of freedom on the event horizon – emphasizing, instead, that we need further assumptions to bridge the derivation of the statistical mechanical entropy in LEQG (which does, at least naively, appear to treat the quantum microstates as located in a thin skin around the event horizon) with the derivation of Hawking radiation (which describes quantum fields external to the horizon).⁷⁷

Similarly, while Wallace argues that we can understand QFT on curved spacetime as an application of the general machinery of modern QFT, Curiel (2023, §6.1) highlights not only a profound difference between the two but also calls attention to the fact that QFT itself (even in the flat spacetime of special relativity) is fraught with difficulties.⁷⁸

While the detection of Hawking radiation may not reveal anything of its fundamental nature, it is nevertheless fair to say that an empirical test of Hawking radiation is required in order to entrench BHT. Yet, the Hawking radiation of typical astrophysical black holes has an extremely weak empirical signature, making its detection a near impossibility. As a black hole's temperature T is proportional to its surface gravity κ – and, thereby, inversely proportional to its mass M – the temperature of the radiation becomes smaller the bigger the black hole considered. Consequently, black holes detectable in astronomical settings would emit radiation of an order several million times smaller than the temperature of the cosmic microwave background (Thébaud, 2019).

⁷⁶ Curiel also notes that this is not standard consilience, since it is not the prediction of the value for an experimentally observed quantity (say), by way of different mechanisms, but rather the prediction of the same (purely theoretical) phenomenon by different theoretical routes.

⁷⁷ Curiel calls this 'the temperature decoupling problem'. Cf. fn. 15.

⁷⁸ See, e.g., Fraser (2016); Kuhlmann (2023).

7.2 Black Hole Information Paradox

If black holes radiate – if they are ordinary statistical mechanical systems obeying standard QM – then we would expect that their radiation and eventual evaporation be described as a fully unitary process, in accordance with QM. Yet, the calculations of the Hawking effect in QFT in curved spacetime describe a process that appears strongly non-unitary. This leads to what is known as the ‘information loss paradox’. In this context, there are two different problems that have been discussed: the ‘final state problem’ (or ‘evaporation time problem’) and the ‘Page time problem’; I briefly introduce each in turn.

Quantum systems are subject to *unitary evolution*: the initial conditions together with the Schrödinger equation fix the future state of the system. This means that unitary evolution ‘preserves information’: if we start with a quantum system in a known state, then quantum theory ensures that the system will evolve in such a way that we can infer the state of the system at some later time, and vice versa – the quantum evolution takes a pure state to a pure state, and this process is reversible. This is challenged by black hole evaporation. If we drop a cup of coffee into the black hole, information about its state still exists (though inaccessible) beyond the horizon, but once the black hole completely evaporates, all that is left is Hawking radiation (as shown in Figure 5 (d)). Hawking radiation is perfectly thermal, meaning that the ‘information’ about what was dropped in cannot be recovered: we cannot infer the previous state of the system given the final state of the Hawking radiation. Another way of seeing the problem is to imagine forming a black hole through the collapse of a pure state (such as a large amplitude gravity wave), and then noting that at the final state of evaporation is a *mixed state*: to an outside observer, the process of black hole formation and evaporation is a pure-to-mixed, irreversible transition. In other words: the evolution is *non-unitary*.

To see why we end up with thermal (mixed state) radiation despite starting from an initial pure state, recall that the effect arises from considering the QFT vacuum in the presence of a horizon. In doing so, we are essentially splitting the original vacuum state into two parts: one interior and one exterior. While the QFT vacuum state is a pure state overall, its degrees of freedom are entangled at short distances, which means that when we consider only half of the space, we find a mixed state on that half. In the heuristic picture described earlier (§7.1) involving pair creation, the two particles are entangled with one another, forming a pure state, but when we consider only the outgoing radiation, we find it in a mixed state looking like a thermal state at the Hawking temperature. [The presentation of the information loss paradox in terms of ‘information’ thus risks being a bit misleading – the basic problem has really to do with the

entanglement, or *correlations*, between matter within the black hole and matter far outside the black hole.]

One difficulty with this story, however, is that the final stages of black hole evaporation occur when the black hole's curvature is of the Planck scale, and so the semiclassical description that gives rise to the 'paradox' is not itself valid here. It cannot tell us what actually happens in the final stages of evaporation. There are various proposals that attempt to exploit this 'loophole' in order to save unitarity: for instance by having the 'information' ejected in the final bursts of energy, or by having it encoded in a non-radiating Planck scale 'remnant' black hole, or something even more exotic (Hawking & Stewart, 1993). None of these proposals appear satisfactory (Marolf, 2017; Unruh & Wald, 2017).

For GR physicists, and most philosophers, this black hole 'final state problem' is not really seen as a paradox – or, indeed, a problem.⁷⁹ On this view, 'the problem' is simply an argument to the conclusion that the black hole evaporation is a non-unitary process (this is often supported, for instance, by claiming that the black hole evaporating spacetime is not globally hyperbolic, and we do not expect unitary evolution in such spacetimes). Early on, Hawking himself (Hawking et al., 1976) stated that the full theory of QG will be one that violates the standard dynamical principle of quantum evolution. Hawking later changed his view, persuaded by arguments in favour of the holographic principle, that 'information' is stored on the event horizon, §7.3. The opposing side in this debate holds that the evolution of an initial pure state to a final mixed state is in conflict with QM. Arguments to this effect are mainly found from within the 'particle physics perspective' on QG (typically, string theorists), which treats unitary evolution as an inviolable principle. Why exactly this need be so, however, is debatable.

The other main form of the 'information loss paradox' is the 'Page time problem'. Here, the problem occurs not in the final stages of evaporation where we expect the semiclassical theory to be invalid anyway, but rather in a regime where the black hole is still big, and where we have no reason to expect the semiclassical description to be incorrect. To understand this paradox, we need to first distinguish between two notions of entropy that we ordinarily use in physics.⁸⁰ The first is the *von Neumann entropy*, S_{vN} , also called the *fine-grained entropy*, or 'entanglement entropy'. Given the density matrix, ρ , describing the quantum state of the system, it is defined as,

⁷⁹ Authors who argue this view include Belot et al. (1999); Maudlin (2017); Unruh & Wald (2017); Wald (2001). Cf. Manchak & Weatherall (2018).

⁸⁰ This presentation follows Almheiri et al. (2021).

$$S_{vN} = -\text{Tr}[\rho \log \rho] \quad (15)$$

This equation quantifies our ignorance about the precise quantum state of the system. It vanishes for a pure state, indicating our complete knowledge of the quantum state. Importantly, it is invariant under unitary time evolution, $\rho \rightarrow U\rho U^{-1}$.

The second notion of entropy is the *coarse-grained entropy*, S_T , also known as the *thermodynamic entropy*. In this case, we have the density matrix ρ , but instead of measuring all observables, we measure only a subset of simple, or coarse-grained, observables, A_i . We then consider all possible density matrices $\tilde{\rho}$ that give the same result as our system for the observables we are tracking, and then compute the von Neumann entropy, $S(\tilde{\rho})$. Finally, we maximize this over all choices of $\tilde{\rho}$. A simple example is the ordinary entropy used in thermodynamics: here, A_i are often chosen to be a few observables, for example, the approximate energy and the volume, and then the thermodynamic entropy is obtained by maximizing the von Neumann entropy among all states with that approximate energy and volume. The coarse-grained entropy obeys the second law of thermodynamics: it tends to increase under unitary time evolution. The generalized entropy, S_{tot} in eq. 14, comprising the entropy of the black hole plus the entropy outside – if it is to be either of these two entropies – would be the coarse grained, thermodynamic entropy (Almheiri et al., 2021).

The von Neumann entropy cannot be larger than the thermodynamic entropy: to see this, realize that we can always consider ρ a candidate $\tilde{\rho}$. Or, in other words, because the thermodynamic entropy provides a measure of the total number of degrees of freedom available to the system, it sets an upper bound on how much the system can be entangled with something else.

$$S_{vN} \geq S_T \quad (16)$$

Now, consider the black hole as an ordinary thermodynamic system. The von Neumann entropy of the matter that formed the black hole will be extremely small compared to the black hole's initial thermodynamic entropy.⁸¹ The system cools through the emission of thermal radiation – that is, through the emission of quanta of radiation in highly mixed states. In the early stages of black hole evaporation, the von Neumann entropy will be almost exactly thermal because the radiation is entangled with the system. [If the original state of

⁸¹ Although it's implausible to think that the stellar precursor to an astrophysical black hole is in a pure state, the thermodynamic entropy of such a precursor is typically negligible compared to the entropy of the black hole that forms from it, so actually the initial purity of the system state plays no essential role here provided that the initial state's von Neumann entropy is much lower than its thermodynamic entropy (Wallace, 2020).

the black hole, considered as a thermodynamic system, is pure, and its dynamics are unitary, then the total state of the system plus radiation is pure. But if the radiation is thermal, this means each emitted particle is in a mixed state, and so must be entangled with some other system (in this case, the black hole) in order for the total state to be pure (no two emitted quanta can be entangled in thermal radiation).] As the system cools, the von Neumann entropy of the radiation rises: the entanglement of the emitted particles increases. However, as the black hole radiates, its area shrinks, and at some point it will not be possible for the emitted radiation to be thermal anymore. This is because the number of degrees of freedom of the black hole is given by its thermodynamic entropy – that is, the area of its horizon. When the area is sufficiently small, it will no longer be possible for the outgoing radiation to be entangled with the system, and at this point, the von Neumann entropy of the radiation will decrease along with the thermodynamic entropy of the black hole.

In summary: if, as in accordance with BHT, black holes are ordinary thermodynamic systems that can be described, by an external observer, as a quantum system with $e^{A/4G}$ microstates, that evolves unitarily, then – from the argument earlier – the von Neumann entropy of the Hawking radiation needs to decrease at some point in order that it not exceed the thermodynamic entropy of the black hole. This argument was made by Page (1993), who suggested that the entropy of the radiation would need to follow the curve (the *Page curve*) indicated in Figure 7, rather than the Hawking curve. The entropy of the radiation at early stages increases until it equals the decreasing thermodynamic entropy of the black hole, and then the two remain equal for the rest of the evaporation process. The time at which the von Neumann entropy of the radiation is maximal (the turnover point) is known as the *Page time*. The Page time for a Schwarzschild black hole is approximately half the total evaporation time: at this point, it is not possible for the black hole's radiation to be thermal, and instead it should be maximally entangled with the early-time radiation.

Now, the problem is that this result is in clear conflict with the predictions regarding Hawking radiation in QFT in curved spacetime. In these calculations, the Hawking radiation is *exactly* thermal, with no entanglement between the early-time and late-time emitted particles. Indeed, there is *no possibility* of such entanglement on this picture, given that each mode of the Hawking radiation is maximally entangled with a partner mode inside the event horizon (a principle known as the *monogamy of entanglement* ensures that these modes cannot be entangled with others), which has no means of escaping.

The Page time problem is indeed a paradox: a clash between the predictions of QFT in curved spacetime and the predictions of black hole statistical mechanics. As stated, it occurs when the black hole is still macroscopic in scale, a

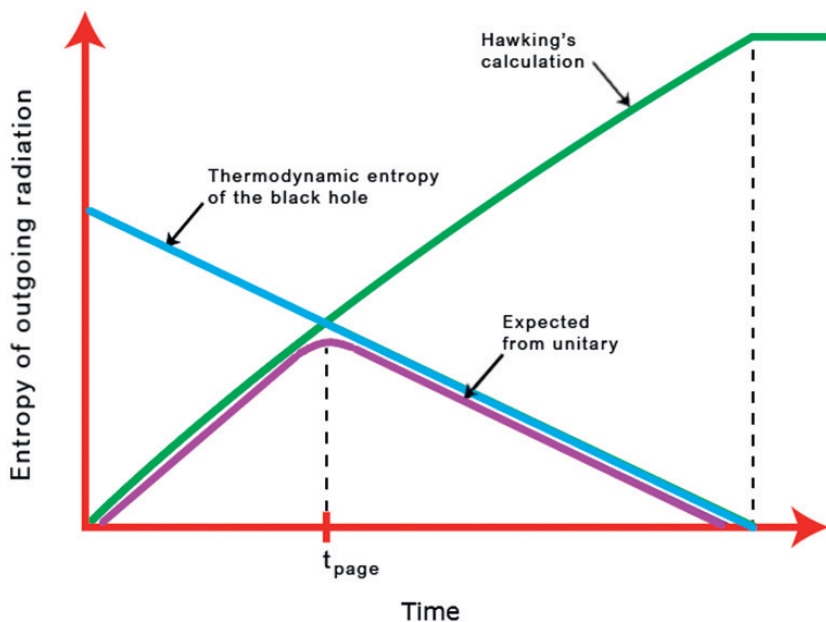


Figure 7 The Page curve: Schematic behaviour of the entropy of the outgoing radiation. In Hawking's calculation remove, the entropy monotonically increases until t_{end} , when the black hole completely evaporates. If the process is unitary, the entropy of radiation must be less than the thermodynamic entropy of the black hole remove, and follows the Page curve ('Expected from unitarity') it saturates this maximum at t_{page} .

regime where *both* these frameworks are expected to be valid, and so closes the 'loophole' left open by the evaporation time problem (where we could attempt to appeal to unknown, exotic QG effects in the Planckian regime). *Prima facie*, there are two ways to resolve the paradox:

1. Accept that QFT in curved spacetime fails as a description of the entire spacetime of an evaporating black hole; retain the statistical-mechanical underpinnings of black hole thermodynamics; try to understand why and when the QFT description breaks down.
2. Retain QFT in curved spacetime, reject black hole statistical mechanics, and find some non-statistical-mechanical understanding of BHT.

Wallace (2020) presents these two options as each packaged with their own explanatory burdens. Option 1, according to Wallace, necessitates explaining when and *why* the QFT description breaks down, *given that* the break-down occurs in regimes 'well within the applicability domain of the theory'. Option 2 necessitates explaining not only *why* we have such a striking theoretical analogy between BHT and ordinary thermodynamic systems, but also *why*

we have managed to reproduce the thermodynamic entropy of black holes in statistical-mechanical calculations in various different contexts. Failure to address the compatibility of this vast network of calculations, he says, would leave these results ‘miraculous’.

Again, though, we might follow Curiel in questioning the necessity of taking on these additional explanatory demands in order to reject one or the other framework, given the tenuous theoretical status of each. In the case of option 1, we might need only to appeal to the fact that ‘semiclassical gravity is an effective field theory with no empirical support – deal with it’. Or, more seriously, it may be philosophically fruitful to question whether, or why, we should be compelled to believe that results developed in the framework of semiclassical gravity are revelatory of genuine insights into a more fundamental theory – given not only the problems associated with semiclassical gravity, but also due to the very nature of effective field theories in general (this is not to imply, of course, that Wallace does not provide sound arguments for his trust in LEQG in various non-Planckian regimes, see Wallace (2022)). In the case of option 2, we might again wish to look more closely at the ‘consilience’ of derivations and explore the inter-theory relations between the various frameworks, and/or bemoan the underavailability of physical assumptions bridging the frameworks generating the statistical-mechanical results with the BHT ones.

Rather than fuel debate here, my goal is to encourage more philosophical interest in exploring these issues. Whichever of these routes one chooses to pursue, it is likely that interesting philosophical insights will be gained – whether concerning QG, BHT, semiclassical gravity, QFT, effective field theory, or more general issues in the philosophy of science such as inter-theory relations, scientific methodology, and scientific progress. In regard to the specific questions being considered in this Element, we can ask how the information loss paradox should be used in motivating and constraining the search for QG. It is a theoretical puzzle that arises by taking seriously our best current theories, developing natural ways of combining them, and then pushing these until breaking point. Surely the paradox serves as an extremely stimulating heuristic and guiding principle, and any approach towards QG that indicates a solution could take this as an indication of pursuit-worthiness, or means of confirmation (given that it is an indication of *consistency* of the approach with low-energy theoretical results). At this stage, however, the resolution of the paradox via new fundamental physics should not be taken as a constraint (criterion of acceptance) on a theory of QG: a *dissolution* is not ruled out, neither a solution ‘at the level of current theories’.

On that note, I close this section by gesturing to the wide variety of different solutions to the paradox being explored in physics. Particularly noteworthy are

recent results involving *replica wormholes*, which show not only that entanglement entropy follows the Page curve,⁸² but also suggest a possible mechanism that allows the ‘information’ to escape via wormholes⁸³ (perhaps as a proxy for nonlocal effects.⁸⁴) Exactly what this could mean physically, and what it could imply for QG, is still very much open to speculation.

7.3 Entropy Bounds

Various theoretical results in BHT suggest that there is an upper bound on how much entropy can be contained in a given region of spacetime. There are several different proposals for the entropy bound, with varying degrees of generality. These results are striking, and in conflict with the picture of the world presented by current theories – the explanation of the entropy bound is motivation for a new, more fundamental, theory.

The idea of an entropy bound originally came from considering the *Geroch process*, which is a mechanism via which the GSL can apparently be violated.⁸⁵ Imagine we prepare a massless box full of energetic radiation with high entropy, far away from the black hole. The mass of the radiation will be attracted by the black hole’s gravitational force. We could then use this weight to drive an engine to produce energy while slowly lowering the box towards the event horizon of the black hole – this process extracts energy (work), but not entropy, from the radiation in the box. If we keep slowly lowering the box towards the event horizon, we can arrange for all the mass-energy of the radiation to have been exhausted when the box reaches the event horizon. Dropping the box into the black hole, then, will not increase the size of the event horizon (because the mass-energy of the black hole does not increase), but will mean that the thermodynamic entropy outside the black hole has decreased.

In response, Bekenstein proposed that there must be a limit to how much entropy can be contained in a given region of spacetime (Bekenstein, 1981). The *Bekenstein bound* states that maximum entropy contained in a spherical system of radius R and total gravitating energy E is $S \leq \frac{2\pi ER}{\hbar}$. For a weakly gravitating system, the Bekenstein bound reduces to $S \leq \frac{A}{4\hbar G}$, where A is the area of a spherical surface fully containing the system.

If we assume, however, that the Bekenstein bound holds also for strongly gravitating systems, we can arrive at a more general conjecture known as the

⁸² This is consistent with other results in the context of AdS/CFT, which has convinced most in the high-energy physics community that unitarity is not violated in black hole evaporation.

⁸³ Almheiri et al. (2020); Penington et al. (2022).

⁸⁴ Marolf & Maxfield (2020).

⁸⁵ Here following Curiel (2023, §5.4.2), but see, e.g., Bekenstein (1972); Bousso (2002) for more details.

spherical entropy bound, using the GSL and considering the *Susskind process*, where, instead of lowering a thermodynamic system into a black hole, we *convert* the system into a black hole (Susskind, 1995). This bound states that the maximum entropy contained in a region of space is $S \leq \frac{A}{4\hbar G}$, where A is the area of any surface bounding the region. To see why this is, imagine that we find a region with entropy greater than $\frac{A}{4\hbar G}$, then we could add energy to it adiabatically without changing its entropy, and thereby transform it into a black hole of area A . Such a black hole would then have entropy greater than the Bekenstein-Hawking entropy, (eq. 13). Thus, if the Bekenstein-Hawking entropy is correct, the entropy of a black hole of surface area A must be the maximum possible entropy of any region of space with surface area A . By definition, then, the entropy of a single Schwarzschild black hole precisely saturates the spherical entropy bound: a black hole is the most entropic object we can put inside a given spherical surface ('t Hooft, 1993). This form of the entropy bound is associated with the *holographic principle*, discussed next.

The heuristic derivation of the spherical entropy bound rests on a large number of fairly strong assumptions (e.g., suitable asymptotic conditions, spherical surface, and gravitational stability of the region). Additionally, the bound suffers several counter examples, so does not have general validity. Bousso (1999) proposed a successful generalization of this bound, known as the *covariant entropy bound*. Heuristically, we can imagine a spherical star with nonzero entropy, S_0 that burns out and undergoes catastrophic gravitational collapse. For an outside observer, the star will form a black hole whose surface area will be at least $4S_0$ in accordance with the GSL. But we can imagine following the star as it falls through its own horizon, and watch it shrink to zero radius as it crushes to a singularity: $A \rightarrow 0$. But, the entropy of the star must still be at least S_0 . This is a case where the spherical entropy bound does not apply, being a regime of dominant gravity. But we can imagine taking a snapshot of the star just before it crunches to a point: light does not have time to cover the entire area of the star before it completely crunches, and so the entropy of the area traversed by the light will always be less than the area of the star's surface. The covariant entropy bound is formulated in terms of *light sheets*, where a light sheet is a way of taking a 'snapshot' of a matter system in light-cone time (so the entropy on a light sheet is given by the entropy of the matter system). More technically, as illustrated in Figure 8, a light sheet associated with a surface is a set of null geodesics leaving the surface orthogonally such that the expansion of the set in the direction going away from the surface is zero or negative; that is, the geodesics are remaining parallel or coming closer together as they

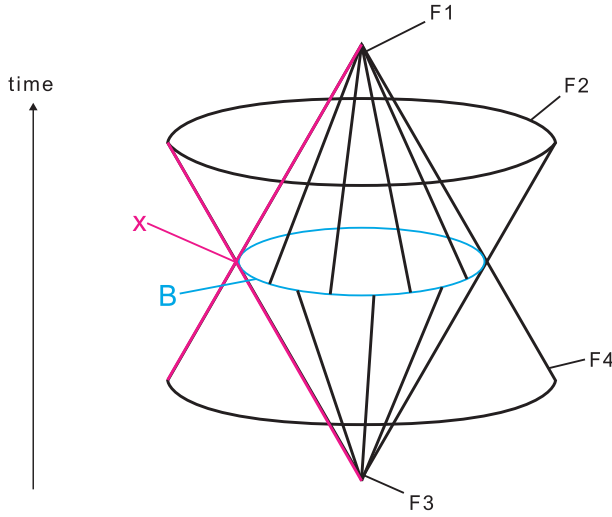


Figure 8 Bousso's light sheet construction, depicting four null hypersurfaces orthogonal to a spherical surface, B (also, the four null geodesics orthogonal to a point x on B , to make the illustration clearer). The two cones F_1, F_3 have negative expansion and hence are light sheets. The covariant entropy bound states that the entropy of each light sheet will not exceed the area of B . The other two families of light rays, F_2, F_4 give the 'skirts'; these have increasing cross-sectional area, and so are not light sheets and their entropy does not relate to the area of B (Bousso, 2002).

get further from the surface. The light sheet continues up until the geodesics intersect or encounter a singularity of spacetime.

The covariant entropy bound simply states that the entropy on a light sheet associated with a surface B will not exceed the area of B , that is, $S \leq A/4G$. This bound appears to hold universally in nature, at least at the semiclassical level – it holds for arbitrary physically realistic matter systems and arbitrary surfaces in any spacetime that is a solution of Einstein's equations. This bound reduces to the spherical entropy bound in special cases where the latter is known to hold. In addition, Bousso (2002) argues that the covariant entropy bound *implies* the GSL, that is, that the GSL can be regarded as a consequence of the covariant entropy bound.

Yet, as Bousso (2002) emphasizes, the origin of the bound remains mysterious: aside from its success, little motivation for a light-like bound can be offered. Although the discovery of the covariant entropy bound was initially inspired by BHT (via the need for a generalization of the Bekenstein bound, which itself was proposed in order to save the GSL), the covariant entropy bound is no longer connected, via derivation, to these discoveries, and stands on its own. Essentially, the principle has outstripped its original motivations.

[Here, it is interesting to recall, as well, that BHT has strange motivations itself. The analogy of the area theorem with the second law of thermodynamics gives perhaps unreasonable weight to the second law itself, which is known not to be fundamental, but open to statistical violations: its appearance is due to perhaps contingent features of our universe.] We cannot look to its mode of discovery in order to justify the covariant entropy bound: instead, its justification is that it holds true in all calculations.

That there could be a maximal bound on entropy inside a region of space-time is itself a striking result that conflicts with our current picture of the world: according to QFT, the number of degrees of freedom in any given region of spacetime is infinite, so this result calls into question the fundamental status of field theory. But the entropy bound is more incredible given its form: it implies that the number of degrees of freedom grows in proportion to the area surrounding the region, rather than the volume of the region. This clashes with the standard dynamics of all other known types of physical systems, where entropy increases as the volume of any spatial region. Notably, even if QG describes a minimal length (e.g., through a discretization of spacetime), which could serve as a short distance cutoff for QFT, the information content of a given region would still be expected to grow with the volume.

As with the other results coming from BHT, however, it is precisely the *inexplicability* of the entropy bound that excites researchers: its origin is expected to be found in a new theory. Because the bound relates aspects of spacetime geometry to the number of quantum states of matter (it involves the fundamental constants of both G and \hbar), the suggestion is that any theory that explains the bound will be a theory of QG: a theory that ‘combines’ GR and QFT. The theory will also need to explain why it is that field theory (QFT) gives the wrong answer. Still, however, it is possible that established physics is correct, and that the entropy bound result is an artefact of the semiclassical approximation. While the entropy bound is promoted as holding true in all calculations, we have to keep in mind that these calculations are still remote from empirical tests.

We can ask how the entropy bound can, or should, be used in the search for QG. Bousso (2002) suggests that the bound may be useful as a *guiding principle* in the context of string theory. This is because the bound holds in situations where string theory suffers limitations, for instance, in highly dynamical geometries, as well as in the non-perturbative regimes of string theory: the hope is that the bound can help aid the development of string theory in these areas. (The AdS/CFT duality is an example of a non-perturbative formulation of string theory where the entropy bound, as well as the holographic principle is satisfied; the idea is to find others, which could physically represent

our universe.) It seems possible that the covariant entropy bound (or the holographic principle) be used as a *postulate* for a candidate theory of QG, similar to how approaches such as causal set theory take *discreteness* as a postulate, for instance. The explanation of the bound can, and has, been appealed to as a *means of confirmation* for QG. Calculations in causal set theory, for instance, give evidence for the existence of an entropy bound which leads to Susskind's spherical entropy bound in the continuum (Rideout & Zohren, 2006). Finally, at this stage, it seems difficult that the explanation of the entropy bound be taken as a *criterion of acceptance* for QG, given its status in semiclassical gravity. The use of the entropy bound in various roles may also be influenced by the attitude we take towards it: Adlam (2023) discusses arguments for and against the epistemic and ontological interpretations of the entropy bound, highlighting the different commitments, challenges, and implications of each.

7.4 Holographic Principle

The most common understanding of the holographic principle is based on the idea that the description of a volume – ‘bulk’ – of space can be equivalently defined on a surface bounding the volume of space. In other words, a physical theory defined only on the $(N-1)$ -dimensional boundary of the region it encloses completely describes the N -dimensional physics of the bulk. The principle is then the claim that the full theory of QG can be reformulated as a theory all of whose degrees of freedom are defined on the boundary – a concrete example is the AdS/CFT duality, mentioned earlier (p. 2.1). The idea, originally from 't Hooft (1993), was popularized, and given a precise interpretation in string theory, by Susskind (1995). The picture is that ‘The three-dimensional world of ordinary experience – the universe filled with galaxies, stars, planets, houses, boulders, and people – is a hologram, an image of reality coded on a distant two-dimensional surface’ (Susskind, 2008, p. 410). Susskind's presentation, following 't Hooft, is based on a response to the information loss paradox, where the suggestion is that the ‘information’ is stored on the black hole horizon.

There are slightly different statements of the holographic principle, and its relation to entropy bounds. Bousso (2002), for instance, takes the holographic principle as simply the claim that the origin of the entropy bound is to be found in a new theory of QG. Smolin (2001), states that, while entropy bounds provide a limit on the number of degrees of freedom in a region of space, the holographic principle is supposed to go further in telling us about the nature of those degrees of freedom, by postulating a form of dynamics in which the quantum evolution of the spacetime and matter fields is described in terms of observables measurable on the screens (surfaces). ‘The holographic principle

is not a relationship between two independent sets of concepts: bulk theories and measures of geometry vs. boundary theories and measures of information. Instead, it is the assertion that in a fundamental theory the first set of concepts must be completely reduced to the second' (Smolin, 2001). These claims suggest, however, that the more fundamental principle is the entropy bound itself, rather than the holographic duality, given that the latter is supposed to explain the former.

Jaksland & Linnemann (2020) usefully distinguish between the holographic duality and entropy bounds, in that the former is *inter-representational*, while the latter are *intra-representational*. The duality is inter-representational because it describes a relationship between different equivalent representations of the same physics; that is, it relates two *separate descriptions* (different theories): one on the 'bulk' and one on the 'boundary'. In the case of the AdS/CFT duality, for instance, the duality relates a string theory in N -dimensional spacetime 'bulk', to a gauge theory (CFT) formulated on the $(N-1)$ -dimensional 'boundary'. On the other hand, the entropy bounds apply only in a single theory, and are thus intra-representational. For this reason (though as should already be evident from the characterization of the holographic principle, earlier), the existence of an entropy bound does not imply the holographic principle. Nevertheless, the holographic principle does suggest an entropy bound in the bulk theory, and its aim is to motivate or explain the existence of the bound. The main question is, then, how the holographic duality can be used to develop a theory of QG.

One suggestion is made by De Haro (2019), which utilizes an account of dualities as isomorphic representations of a single 'common-core' theory.⁸⁶ We can then distinguish between different functions of dualities in theory-construction. The *theoretical function* of dualities is their use in developing the common-core theory that is represented on both sides of the duality: the idea is to utilize the dual theories, and exploit the relationship between them, to extract the content of this common physics that is being represented differently by each. The *heuristic function*, rather than discovering and describing the 'equivalent physics' (the existing physics of the underlying common-core theory), aims at discovering *new physics* beyond the common-core theory: constructing a new theory that supersedes the common-core theory. While the theoretical function relies on the exactness of the duality, and it being a non-perturbative description of equivalent physics (in order to identify the physics of the common-core), the heuristic function requires breaking the (exactness of the) duality, and showing that it is recovered only approximately (e.g., in some

⁸⁶ Cf. De Haro & Butterfield (2021); Rickles (2017).

appropriate limit, or perturbatively) from a more fundamental theory beyond the common-core theory. The heuristic function thus treats the duality as an ‘intermediate step’ in discovering new physics.

Jaksland & Linnemann (2020) argue that, given that holography is inter-representational, it cannot, as it stands, tell us anything ‘outside’ of itself (it only tells us about the relationship between features on the two different sides of the duality). In order for the duality to be useful for discovering new physics, we need to render it intra-representational – so that we are working within a single description, which, they argue, need not be the common-core theory. The heuristic use of the duality, for Jaksland & Linnemann (2020), is to get at the intra-representational description. On this account, it does not make sense to recover the duality, whether approximately or not. As a simple example of the heuristic impotency of duality, take the *Ryu-Takayanagi formula* (Ryu & Takayanagi, 2006). If we consider a spatial slice, Σ of an AdS spacetime on whose boundary we define the dual CFT, the formula is a conjecture relating the entanglement entropy S of the CFT (boundary theory) in some spatial subregion, with the area A of a ‘Ryu-Takayanagi surface’ in the bulk theory, as $S = A/(4G)$. This formula is clearly inter-representational. As such, Jaksland & Linnemann (2020) argue that it cannot testify to the origin of the area in terms of entropy or vice versa: in the representation where the area exists, the entropy is not there, and where the entropy exists, the area is not there.

So, the holographic duality can be used as a *guiding principle* in the search for QG, but this involves either the *inexactness* of the duality, following De Haro (2019), or its transformation into a non-dual description, according to Jaksland & Linnemann (2020). The holographic principle can also be used as a *postulate*, as for instance Smolin (2017), who presents it as one of four principles of QG, in the attempt to build QG as a principle theory, in Einstein’s sense. The holographic principle could potentially function as a constitutive principle – indeed, it seems the most likely candidate for a coordinating principle defining a new physical framework (Crowther, 2021). For instance, Bousso (2002, p. 861) suggests that the holographic principle may not only aid the search for a non-perturbative version of string theory, but ‘could also contribute to a background-independent formulation that would illuminate the conceptual foundation of string theory’. Smolin (2001) also presents his weak holographic principle as playing something like a coordinating role. The holographic principle can also be appealed to as a *means of confirmation*: for instance the AdS/CFT duality as embodying the holographic principle is presented as an argument in favour of string theory, especially in its

background independent, non-perturbative form. Additionally, if we follow De Haro (2019), the approximate recovery of the duality from a more-fundamental theory of QG (a theory of new physics, beyond the common-core theory embodied by the duality), could serve as a means of confirmation (though Jaksland & Linnemann (2020) argue against this). Finally, considering taking the holographic principle as a *criterion of acceptance* for QG requires further argument. This may be prompted, perhaps, if the entropy bound is taken as a fundamental result motivating QG – the holographic principle would then be required in order to explain the origin of the entropy bound.

References

- Adlam, E. (2022). Tabletop experiments for quantum gravity are also tests of the interpretation of quantum mechanics. *Foundations of Physics*, 52(115), 1–43.
- Adlam, E. (2023). Are entropy bounds epistemic? *arXiv:2303.10781v3 [physics.hist-ph]*.
- Aharony, O., Marsano, J., Minwalla, S., Papadodimas, K., & Van Raamsdonk, M. (2004). The Hagedorn - deconfinement phase transition in weakly coupled large N gauge theories. *Advances in Theoretical and Mathematical Physics*, 8, 603–696.
- Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., & Tajdini, A. (2020). Replica wormholes and the entropy of hawking radiation. *Journal of High Energy Physics*, 2020(5), 1–42.
- Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., & Tajdini, A. (2021). The entropy of hawking radiation. *Reviews of Modern Physics*, 93, 035002.
- Ambjørn, J., Jurkiewicz, J., & Loll, R. (2006). The universe from scratch. *Contemporary Physics*, 47(2), 103–117.
- Ambjørn, J., Jurkiewicz, J., & Loll, R. (2009). Quantum gravity: The art of building spacetime. In D. Oriti (Ed.), *Approaches to Quantum Gravity* (pp. 341–359). Cambridge: Cambridge University Press.
- Anderson, E. (2017). *The Problem of Time: Quantum Mechanics versus Special Relativity*. Cham: Springer.
- Ashtekar, A. (1995). Mathematical problems of non-perturbative quantum general relativity. In B. Julia & J. Zinn-Justin (Eds.), *Les Houches, Session LVII, 1992: Gravitation and Quantizations* (pp. 181–283). Amsterdam: Elsevier Science.
- Ashtekar, A. (2005). The winding road to quantum gravity. *Current Science*, 89, 2064–2074.
- Ashtekar, A. & Geroch, R. (1974). Quantum theory and gravitation. *Reports on Progress in Physics*, 37, 1211–1256.
- Bain, J. (2008). Condensed matter physics and the nature of spacetime. In D. Dieks (Ed.), *The Ontology of Spacetime II* (pp. 301–329). Oxford: Elsevier.
- Barbado, L. C., Barceló, C., Garay, L. J., & Jannes, G. (2016). Hawking versus Unruh effects, or the difficulty of slowly crossing a black hole horizon. *Journal of High Energy Physics*, 2016(10), 161–174.

- Barbour, J. (2000). *The End of Time*. Oxford: Oxford University Press.
- Barceló, C., Liberati, S., & Visser, M. (2011). Analogue gravity. *Living Reviews in Relativity*, 14(3), 1–159.
- Barceló, C., Visser, M., & Liberati, S. (2001). Einstein gravity as an emergent phenomenon? *International Journal of Modern Physics D*, 10(6), 799–806.
- Batterman, R. W. (2018). Autonomy of theories: An explanatory problem. *Noûs*, 52, 858–873.
- Batterman, R. W. (2011). Emergence, singularities, and symmetry breaking. *Foundations of Physics*, 41, 1031–1050.
- Bekenstein, J. (1981). A universal upper bound on the entropy to energy ratio for bounded systems. *Physical Review D*, 23(2), 287.
- Bekenstein, J. D. (1972). Black holes and the second law. *Lettere al Nuovo Cimento*, 4, 737–740.
- Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D*, 7(8), 2333–2346.
- Bekenstein, J. D. (1974). Generalized second law of thermodynamics in black-hole physics. *Physical Review D*, 9, 3292–3300.
- Belot, G. (2011). Background-independence. *General Relativity and Gravitation*, 43(10), 2865–2884.
- Belot, G., Earman, J., & Ruetsche, L. (1999). The hawking information loss paradox: The anatomy of a controversy. *British Journal for the Philosophy of Science*, 50(2), 189–229.
- Bern, Z. (2002). Perturbative quantum gravity and its relation to gauge theory. *Living Reviews in Relativity*, 5(5), 1–57.
- Bohr, N. & Rosenfeld, L. (1933). Zur frage der messbarkeit der elektromagnetischen feldgrössern. *Det Kgl. Danske Videnskabernes Selskab. Matematisk-fysiske Meddelelser*, XII(8), 1–33.
- Bonanno, A., Eichhorn, A., Gies, H., et al. (2020). Critical reflections on asymptotically safe gravity. *Frontiers in Physics*, 8, 1–30.
- Bose, S., Mazumdar, A., Morley, G. W., et al. (2017). Spin entanglement witness for quantum gravity. *Physical Review Letters*, 119, 240401.
- Bousso, R. (1999). A covariant entropy conjecture. *Journal of High Energy Physics*, 7.
- Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74, 825–874.
- Brandenberger, R., Mukhanov, V., & Sornborger, A. (1993). Cosmological theory without singularities. *Physical Review D*, 48, 1629–1642.
- Brown, H. (2005). *Physical Relativity: Space-Time Structure from a Dynamical Perspective*. Oxford: Oxford University Press.

- Burgess, C. P. (2004). Quantum gravity in everyday life: General relativity as an effective field theory. *Living Reviews in Relativity*, 7(5), 1–56.
- Butterfield, J. & Bouatta, N. (2015). Renormalization for philosophers. In T. Bigaj & C. Wüthrich (Eds.), *Metaphysics in Contemporary Physics*, Poznan Studies in Philosophy of Science (pp. 437–485). Boston: Brill
- Butterfield, J. & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The Arguments of Time* (pp. 116–168). Oxford: Oxford University Press.
- Butterfield, J. & Isham, C. (2001). Spacetime and the philosophical challenge of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics Meets Philosophy at the Planck Scale* (pp. 33–89). Cambridge: Cambridge University Press.
- Calmet, X. & Latosh, B. (1998). Dark matter in quantum gravity. *European Physical Journal C*, 78(520), 1–4.
- Cao, T. Y. & Schweber, S. S. (1993). The conceptual foundations and the philosophical aspects of renormalization theory. *Synthese*, 97(1), 33–108.
- Caravelli, F. & Markopoulou, F. (2011). Properties of quantum graphity at low temperature. *Physical Review D*, 84(2), 024002.
- Carlip, S. (2001). Quantum gravity: A progress report. *Reports on Progress in Physics*, 64(8), 885.
- Chiribella, G. (2020). Quantum superpositions of causal structures. *Critical Hermeneutics*, 4(special II), 1–24.
- Colella, R., Overhauser, A. W., & Werner, S. A. (1975). Observation of gravitationally induced quantum interference. *Physical Review Letters*, 34, 1472–1474.
- Cotler, J. & Strominger, A. (2022). The universe as a quantum encoder. *arXiv:2201.11658 [hep-th]*.
- Crowther, K. (2015). Decoupling emergence and reduction in physics. *European Journal for Philosophy of Science*, 5(3), 419–445.
- Crowther, K. (2016). *Effective Spacetime: Understanding Emergence in Effective Field Theory and Quantum Gravity*. Heidelberg: Springer.
- Crowther, K. (2018). Inter-theory relations in quantum gravity: Correspondence, reduction, and emergence. *Studies in History and Philosophy of Modern Physics*, 63, 74–85.
- Crowther, K. (2019). When do we stop digging? Conditions on a fundamental theory of physics. In A. Aguirre, B. Foster, & Z. Merali (Eds.), *What Is 'Fundamental'?* Cham: Springer, 123–133.
- Crowther, K. (2020). What is the point of reduction in science? *Erkenntnis*, 85(6), 1437–1460.

- Crowther, K. (2021). Defining a crisis: The roles of principles in the search for a theory of quantum gravity. *Synthese*, 198(Suppl 14), 3489–3516.
- Crowther, K. & De Haro, S. (2022). Four attitudes towards singularities in the search for a theory of quantum gravity. In A. Vassallo (Ed.), *The Foundations of Spacetime Physics: Philosophical Perspectives* (pp. 223–250). London: Routledge.
- Crowther, K. & Linnemann, N. (2019). Renormalizability, fundamentality, and a final theory: The role of UV-completion in the search for quantum gravity. *British Journal for the Philosophy of Science*, 70(2), 377–406.
- Crowther, K., Linnemann, N. S., & Wüthrich, C. (2021). What we cannot learn from analogue experiments. *Synthese*, 198, 3701–3726.
- Curiel, E. (1999). The analysis of singular spacetimes. *Philosophy of Science*, 66(3), 145.
- Curiel, E. (2023). Singularities and black holes. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/sum2023/entries/spacetime-singularities/>.
- Dafermos, M. & Luk, J. (2017). The interior of dynamical vacuum black holes I: The c° -stability of the Kerr Cauchy horizon. *arXiv:1710.01722*.
- Dardashti, R., Hartmann, S., Thébault, K., & Winsberg, E. (2019). Hawking radiation and analogue experiments: A Bayesian analysis. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67, 1–11.
- Dardashti, R., Thébault, K., & Winsberg, E. (2017). Confirmation via analogue simulation: What dumb holes could tell us about gravity. *British Journal for the Philosophy of Science*, 1(68), 55–89.
- Dawid, R. (2013). *String Theory and the Scientific Method*. Cambridge: Cambridge University Press.
- De Haro, S. (2019). The heuristic function of duality. *Synthese*, 196, 5169–5203. <https://doi.org/10.1007/s11229-018-1708-9>.
- De Haro, S. & Butterfield, J. (2021). On symmetry and duality. *Synthese*, 198(4), 2973–3013.
- De Haro, S., van Dongen, J., Visser, M., & Butterfield, J. (2020). Conceptual analysis of black hole entropy in string theory. *Studies in History and Philosophy of Modern Physics*, 69, 82–111.
- Dewitt, C. M. & Rickles, D. (2011). *The role of gravitation in physics: Report from the 1957 Chapel Hill Conference*.
- Doboszewski, J. (2019). Relativistic spacetimes and definitions of determinism. *European Journal for Philosophy of Science*, 9(2), 24.
- Doboszewski, J. (2020). Epistemic holes and determinism in classical general relativity. *British Journal for the Philosophy of Science*, 71(3), 1093–1111.

- Donoghue, J. F. (2020). A critique of the asymptotic safety program. *Frontiers in Physics*, 8, 56.
- Dowker, F. (2020). Being and becoming on the road to quantum gravity: Or, the birth of a baby is not a baby. In N. Huggett, K. Matsubara, & C. Wüthrich (Eds.), *Beyond Spacetime: The Foundations of Quantum Gravity* (pp. 133–142). Cambridge: Cambridge University Press.
- Dvali, G., Giudice, G. F., Gomez, C., & Kehagias, A. (2011). UV-completion by classicalization. *Journal of High Energy Physics*, 2011(8), 1–31.
- Dyson, F. W., Eddington, A. S., & Davidson, C. (1920). A determination of the deflection of light by the sun's gravitational field, from observations made at the total eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, 220, 291–333.
- Earman, J. (1992). Cosmic censorship. *PSA: Proceedings*, 2, 171–180.
- Earman, J. (1995). *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. New York: Oxford University Press.
- Earman, J. (1996). Tolerance for spacetime singularities. *Foundations of Physics*, 26(5), 623–640.
- Earman, J. (2006). The implications of general covariance for the ontology and ideology of spacetime. In D. Dieks (Ed.), *The Ontology of Spacetime* (pp. 3–23). Amsterdam: Elsevier.
- Eichhorn, A. (2019). An asymptotically safe guide to quantum gravity and matter. *Frontiers in Astronomy and Space Sciences*, 5, 47.
- Ellis, G. F., Meissner, K., & Nicolai, H. (2018). The physics of infinity. *Nature Physics*, 14, 770–772.
- Eppley, K. & Hannah, E. (1977). The necessity of quantizing the gravitational field. *Foundations of Physics*, 7(1), 51–68.
- Fraser, D. (2011). How to take particle physics seriously: A further defence of axiomatic quantum field theory. *Studies in History and Philosophy of Modern Physics*, 42(2), 126–135.
- Fraser, J. D. (2016). *What is quantum field theory? Idealisation, explanation and realism in high energy physics*. PhD thesis, University of Leeds.
- Fraser, J. D. (2020). The real problem with perturbative quantum field theory. *British Journal for the Philosophy of Science*, 71(2), 391–413.
- Gies, H. & Jaeckel, J. (2004). Renormalization flow of QED. *Physical Review Letters*, 93(11), 110405.
- Giulini, D. (2007). Remarks on the notions of general covariance and background independence. In I.-O. Stamatescu & E. Seiler (Eds.), *Lecture Notes in Physics*, volume 721 (pp. 105–120). Berlin: Springer.

- Gomes, H. (2020). Back to parmenides. In N. Huggett, K. Matsubara, & C. Wüthrich (Eds.), *Beyond Spacetime: The Foundations of Quantum Gravity* (pp. 176–206). Cambridge: Cambridge University Press.
- Goswami, K., Cao, Y., Paz-Silva, G. A., Romero, J., & White, A. G. (2020). Increasing communication capacity via superposition of order. *Physical Review Research*, 2, 033292.
- Großardt, A. (2021). Comment on ‘do gedankenexperiments compel quantization of gravity’. *arXiv:2107.14666 [gr-qc]*.
- Großardt, A. (2022). Three little paradoxes: Making sense of semiclassical gravity. *AVS Quantum Science*, 4(1), 010502.
- Hagar, A. (2014). Discrete or Continuous? The Quest for Fundamental Length in Modern Physics. Cambridge: Cambridge University Press.
- Harlow, D. (2016). Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics*, 88(1), 015002.
- Hartmann, S. (2002). On correspondence. *Studies in History and Philosophy of Modern Physics*, 33(1), 79–94.
- Hawking, S. (1971). Gravitational radiation from colliding black holes. *Physical Review Letters*, 26, 1344.
- Hawking, S. (1974). Black hole explosions? *Nature*, 248, 30–31.
- Hawking, S. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, 43, 199–220.
- Hawking, S., King, A., & McCarthy, P. (1976). A new topology for curved space-time which incorporates the causal, differential, and conformal structures. *Journal of Mathematical Physics*, 17(2), 174–181.
- Hawking, S. & Stewart, J. (1993). Naked and thunderbolt singularities in black hole evaporation. *Nuclear Physics B*, 400(1–3), 393–415.
- Hawking, S. W. & Ellis, G. F. (1973). *The Large-Scale Structure of Space-Time*. Cambridge: Cambridge University Press.
- Held, A. (2019). *From particle physics to black holes: The predictive power of asymptotic safety*. PhD thesis, University of Heidelberg, Heidelberg.
- Henson, J. (2009). The causal set approach to quantum gravity. In D. Oriti (Ed.), *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter* (pp. 393–413). Cambridge: Cambridge University Press.
- Hossenfelder, S. (2013). Minimal length scale scenarios for quantum gravity. *Living Reviews in Relativity*, 16, 1–90 (2).
- Hu, B.-L. (2009). Emergent/quantum gravity: macro/micro structures of space-time. In H. T. Elze, L. Diosi, L. Fronzoni, J. Halliwell, & G. Vitiello (Eds.), *Fourth International Workshop Dice 2008: From Quantum Mechanics through Complexity to Spacetime: The Role of Emergent Dynamical*

- Structures*, volume 174 of *Journal of Physics Conference Series* (pp. 12015–12015). Bristol: IOP.
- Huggett, N. & Callender, C. (2001). Why quantize gravity (or any other field for that matter)? *Philosophy of Science*, 68(3), S382–S394.
- Huggett, N., Linnemann, N., & Schneider, M. D. (2023). *Quantum Gravity in a Laboratory?* Elements in the Foundations of Contemporary Physics. Cambridge: Cambridge University Press.
- Huggett, N. & Vistarini, T. (2015). Deriving general relativity from string theory. *Philosophy of Science*, 82(5), 1163–1174.
- Huggett, N. & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics*, 44(3), 276–285.
- Isham, C. (1993). Canonical quantum gravity and the problem of time. In L. Ibort & M. Rodriguez (Eds.), *Integrable Systems, Quantum Groups, and Quantum Field Theories* (pp. 157–288). Dordrecht: Kluwer.
- Jackiw, R. (1999). The unreasonable effectiveness of quantum field theory. In T. Cao (Ed.), *Conceptual Foundations of Quantum Field Theory* (pp. 148–159). Cambridge: Cambridge University Press.
- Jackiw, R. (2000). What good are quantum field theory infinities? In A. Fokas, A. Grigoryan, T. Kibble, & B. Zegarlinski (Eds.), *Mathematical Physics 2000* (pp. 101–110). Singapore: World Scientific.
- Jacobson, T. (1995). Thermodynamics of spacetime: The Einstein equation of state. *Physical Review Letters*, 75, 1260–1263.
- Jakobsen, R. & Linnemann, N. S. (2020). Holography without holography: How to turn inter-representational into intra-theoretical relations in ads/cft. *Studies in History and Philosophy of Modern Physics*, 71, 101–117.
- Kao, M. (2019). Unification beyond justification: A strategy for theory development. *Synthese*, 196(8), 3263–3278.
- Kaplan, D. B. (2005). Five lectures on effective field theory. *arXiv:nucl-th/0510023*.
- Kastner, R. E. & Kauffman, S. (2018). Are dark energy and dark matter different aspects of the same physical process? *Frontiers in Physics*, 6(71), 1–4.
- Kent, A. (2018). Simple refutation of the Eppley-Hannah argument. *Classical and Quantum Gravity*, 35(24), 245008.
- Kiefer, C. (2007a). *Quantum Gravity* (2nd ed.). Oxford: Oxford University Press.
- Kiefer, C. (2007b). Why quantum gravity? In I.-O. Stamatescu & E. Seiler (Eds.), *Approaches to Fundamental Physics*, volume 721 of *Lecture Notes in Physics* (pp. 123–130). Berlin: Springer.

- Kiefer, C. (2013). Conceptual problems in quantum gravity and quantum cosmology. *ISRN Mathematical Physics*, 2013, 509316.
- Kretschmann, E. (1917). Über den physikalischen sinn der relativitätspostulate: A. einsteins neue und seine ursprüngliche relativitätstheorie. *Annalen der Physik*, 53, 575–614.
- Kuchař, K. (1999). The problem of time in quantum geometrodynamics. In J. Butterfield (Ed.), *The Arguments of Time* (pp. 169–196). Oxford: Oxford University Press.
- Kuhlmann, M. (2023). Quantum field theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Online: <https://plato.stanford.edu/archives/sum2023/entries/quantum-field-theory/>
- Landau, L., Abrikosov, A., & Khalatnikov, I. (1954). The removal of infinities in quantum electrodynamics. *Doklady Akademii Nauk SSSR*, 1–4.
- Liberati, S., Girelli, F., & Sindoni, L. (2009). Analogue models for emergent gravity. Online: <https://arxiv.org/abs/0909.3834>.
- Linnemann, N. S. & Visser, M. R. (2018). Hints towards the emergent nature of gravity. *Studies in History and Philosophy of Modern Physics*, 64, 1 – 13.
- Loll, R. (1998). Discrete approaches to quantum gravity in four dimensions. *Living Reviews in Relativity* 1(13), 1–53.
- Loll, R. (2019). Quantum gravity from causal dynamical triangulations: A review. *Classical and Quantum Gravity*, 37(1), 013002.
- Loll, R., Fabiano, G., Frattulillo, D., & Wagner, F. (2022). Quantum gravity in 30 questions. *arXiv:2206.06762 [hep-th]*.
- Lüscher, M. & Weisz, P. (1987). Scaling laws and triviality bounds in the lattice ϕ^4 theory: (I). one-component model in the symmetric phase. *Nuclear Physics B*, 290, 25–60.
- Maldacena, J. (1998). The large N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, 2, 231–252.
- Manchak, J. B. & Weatherall, J. O. (2018). Paradox regained? A brief comment on maudlin on black hole information loss. *Foundations of Physics*, 48(6), 611–627.
- Manohar, A. (1997). Effective field theories. In H. Latal & W. Schweiger (Eds.), *Perturbative and Nonperturbative Aspects of Quantum Field Theory*, volume 479 of *Lecture Notes in Physics* (pp. 311–362). Berlin: Springer.
- Markopoulou, F. (2009a). New directions in background independent quantum gravity. In D. Oriti (Ed.), *Approaches to Quantum Gravity* (pp. 129–149). Cambridge: Cambridge University Press.
- Markopoulou, F. (2009b). Space does not exist, so time can. *arXiv:0909.1861 [gr-qc]*, 1–9.

- Marletto, C. & Vedral, V. (2017). Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity. *Physical Review Letters*, 119, 240402.
- Marolf, D. (2017). The black hole information problem: Past, present, and future. *Reports on Progress in Physics*, 80(9), 092001.
- Marolf, D. & Maxfield, H. (2020). Transcending the ensemble: Baby universes, spacetime wormholes, and the order and disorder of black hole information. *Journal of High Energy Physics*, 2020(8), 1–72.
- Mattingly, J. (2005). Is quantum gravity necessary? In A. Kox & J. Eisenstaedt (Eds.), *The Universe of General Relativity* (pp. 327–338). Basel: Birkhäuser.
- Mattingly, J. (2006). Why Eppley and Hannah's thought experiment fails. *Physical Review D*, 73, 062025.
- Mattingly, J. (2009). Mongrel gravity. *Erkenntnis*, 70(3), 379–395.
- Maudlin, T. (1996). On the unification of physics. *Journal of Philosophy*, 93(3), 129–144.
- Maudlin, T. (2012). *Philosophy of Physics: Space and Time*. Princeton: Princeton University Press.
- Maudlin, T. (2017). (Information) Paradox lost. Online: <https://arxiv.org/abs/1705.03541>.
- Meissner, K. A. (2004). Black-hole entropy in loop quantum gravity. *Classical and Quantum Gravity*, 21(22), 5245–5251.
- Misner, C. (1969). Absolute zero of time. *Physical Review*, 186(5), 1328–1333.
- Misner, C. W., Thorne, K. S., & Wheeler, J. A. (2017). *Gravitation*. Princeton: Princeton University Press.
- Morganti, M. (2020). Fundamentality in metaphysics and the philosophy of physics. Part ii: The philosophy of physics. *Philosophy Compass*, 15(10), 1–14.
- Morrison, M. (2000). *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*. Cambridge: Cambridge University Press.
- Myrvold, W. C. (2003). A Bayesian account of the virtue of unification. *Philosophy of Science*, 70(2), pp. 399–423.
- Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy*, 70(7), 181–201.
- Niedermaier, M. & Reuter, M. (2006). The asymptotic safety scenario in quantum gravity. *Living Reviews in Relativity*, 9(5), 1–173, www.livingreviews.org/lrr-2006-5.
- Norton, J. (2003). General covariance, gauge theories and the Kretschmann objection. In K. Brading & E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections* (pp. 110–123). Cambridge: Cambridge University Press.

- Oppenheim, J. (2023). A postquantum theory of classical gravity? *Physical Review X*, 13, 041040.
- Oriti, D. (2009). The group field theory approach to quantum gravity. In D. Oriti (Ed.), *Approaches to Quantum Gravity: Toward a New Understanding of Space Time and Matter* (pp. 310–331). Cambridge: Cambridge University Press.
- Oriti, D. & Pang, X. (2021). Phantom-like dark energy from quantum gravity. *Journal of Cosmology and Astroparticle Physics*, 2021(12), 040.
- Overhauser, A. W. & Colella, R. (1974). Experimental test of gravitationally induced quantum interference. *Physical Review Letters*, 33, 1237–1239.
- Padmanabhan, T. (2004). Equipartition of energy in the horizon degrees of freedom and the emergence of gravity. *Classical and Quantum Gravity*, 21, 1129–1136.
- Padmanabhan, T. (2010). Thermodynamical aspects of gravity: New insights. *Reports on Progress in Physics*, 73(4), 046901.
- Page, D. N. (1993). Information in black hole radiation. *Physical Review Letters*, 71, 3743–3746.
- Page, D. N. & Geilker, C. D. (1981). Indirect evidence for quantum gravity. *Physical Review Letters*, 47, 979–982.
- Penington, G., Shenker, S. H., Stanford, D., & Yang, Z. (2022). Replica wormholes and the black hole interior. *Journal of High Energy Physics* 2022(205). 1–87.
- Penrose, R. (1965). Gravitational collapse and space-time singularities. *Physical Review Letters*, 14, 57–59.
- Penrose, R. (1979). Singularities and time-asymmetry. In S. Hawking & W. Israel (Eds.), *General Relativity: An Einstein Centenary Survey* (pp. 581–638). Cambridge: Cambridge University Press.
- Penrose, R. (2002). Gravitational collapse: The role of general relativity. *General Relativity and Gravitation*, 34(7), 1141–1165.
- Penrose, R. (2004). *The Road to Reality*. London Jonathan Cape.
- Penrose, R. (2014). On the gravitization of quantum mechanics 1: Quantum state reduction. *Foundations of Physics*, 44, 557–575.
- Peres, A. & Terno, D. R. (2001). Hybrid classical-quantum dynamics. *Physical Review A*, 63, 022101.
- Pitts, J. B. (2006). Absolute objects and counterexamples: Jones-Geroch dust, Torretti constant curvature, tetrad-spinor, and scalar density. *Studies in History and Philosophy of Modern Physics*, 37, 347–371.
- Polchinski, J. (2017). Dualities of fields and strings. *Studies in History and Philosophy of Modern Physics*, 59, 6–20.

- Pooley, O. (2010). Substantive general covariance: Another decade of dispute. In M. Su'arez, M. Dorato, & M. R'edei (Eds.), *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association* (pp. 197–210). Cham: Springer.
- Pooley, O. (2017). Background independence, diffeomorphism invariance, and the meaning of coordinates. In D. Lehmkuhl, G. Schieman, & E. Scholz (Eds.), *Towards a Theory of Spacetime Theories* (pp. 105–144). Basel: Birkhäuser.
- Post, H. (1971). Correspondence, invariance and heuristics: In praise of conservative induction. *Studies in History and Philosophy of Science Part A*, 2(3), 213–255.
- Radder, H. (1991). Heuristics and the generalized correspondence principle. *British Journal for the Philosophy of Science*, 42, 195–226.
- Read, J. (2023). *Background Independence in Classical and Quantum Gravity*. Oxford: Oxford University Press.
- Rickles, D. (2006). Time and structure in canonical gravity. In D. Rickles, S. French, & J. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity* (pp. 152–196). Oxford: Oxford University Press.
- Rickles, D. (2008). Who's afraid of background independence? In D. Deiks (Ed.), *The Ontology of Spacetime II* (pp. 133–152). Amsterdam: Elsevier.
- Rickles, D. (2017). Dual theories: Same but different or different but same? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 59, 61–67.
- Rideout, D. & Zohren, S. (2006). Evidence for an entropy bound from fundamentally discrete gravity. *Classical and Quantum Gravity*, 23(22), 6195.
- Rosenfeld, L. (1963). On quantization of fields. *Nuclear Physics*, 40, 353–356.
- Rovelli, C. (1996). Black hole entropy from loop quantum gravity. *Physical Review Letters*, 77(16), 3288–3291.
- Rovelli, C. (2000). Notes for a brief history of quantum gravity. In V. G. Gurzadyan, R. T. Jantzen, & R. Ruffini (Eds.), *9th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (MG 9)* (pp. 742–768).
- Rovelli, C. (2001). Quantum spacetime: What do we know? In C. Callender & N. Huggett (Eds.), *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity* (pp. 101–124). Cambridge: Cambridge University Press.
- Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

- Rovelli, C. (2020). Space and time in loop quantum gravity. In N. Huggett, K. Matsubara, & C. Wüthrich (Eds.), *Beyond Spacetime: The Foundations of Quantum Gravity* (pp. 117–132). Cambridge: Cambridge University Press.
- Rovelli, C. & Vidotto, F. (2014). *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory*. Cambridge: Cambridge University Press.
- Rydving, E., Aurell, E., & Pikovski, I. (2021). Do Gedanken experiments compel quantization of gravity? *Physical Review D*, 104, 086024.
- Ryu, S. & Takayanagi, T. (2006). Holographic derivation of entanglement entropy from the anti-de sitter space/conformal field theory correspondence. *Physical Review Letters*, 96, 181602.
- Sakharov, A. (1967). Vacuum quantum fluctuations in curved space and the theory of gravitation. *Doklady Akademii Nauk SSSR*, 177, 70–71.
- Sakharov, A. D. (2000). Vacuum quantum fluctuations in curved space and the theory of gravitation. *General Relativity and Gravitation*, 32, 365–367.
- Salimkhani, K. (2018). Quantum gravity: A dogma of unification? In A. Christian, D. Hommen, N. Retzlaff, & G. Schurz (Eds.), *Philosophy of Science: European Studies in Philosophy of Science, Vol 9*. (pp. 23–41). Cham: Springer.
- Salimkhani, K. (2021). Explaining unification in physics internally. *Synthese*, 198(6), 5861–5882.
- Schindler, S. (2018). *Theoretical Virtues in Science: Uncovering Reality through Theory*. Cambridge: Cambridge University Press.
- Smeenk, C. (2013). Philosophy of cosmology. In R. Batterman (Ed.), *Oxford Handbook of Philosophy of Physics* (pp. 607–652). Oxford: Oxford University Press.
- Smeenk, C. & Wüthrich, C. (2021). Determinism and general relativity. *Philosophy of Science*, 88(4), 638–664.
- Smolin, L. (2001). The strong and weak holographic principles. *Nuclear Physics B*, 601(1), 209–247.
- Smolin, L. (2006). The case for background independence. In D. Rickles, S. French, & J. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity* (pp. 196–239). Oxford: Oxford University Press.
- Smolin, L. (2013). *Time Reborn: From the Crisis of Physics to the Future of the Universe*. London: Allen Lane.
- Smolin, L. (2017). Four principles for quantum gravity. In J. Bagla & S. Engineer (Eds.), *Gravity and the Quantum*, volume 187 of *Fundamental Theories of Physics* (pp. 427–450). Cham: Springer.

- Smolin, L. (2020). Temporal relationalism. In N. Huggett, K. Matsubara, & C. Wüthrich (Eds.), *Beyond Spacetime: The Foundations of Quantum Gravity* (pp. 143–175). Cambridge: Cambridge University Press.
- Stachel, J. (2006). Structure, individuality and quantum gravity. In D. Rickles & S. French (Eds.), *The Structural Foundations of Quantum Gravity* (pp. 53–82). Oxford: Oxford University Press.
- Steinhauer, J. (2016). Observation of quantum Hawking radiation and its entanglement in an analogue black hole. *Nature Physics*, 12(10), 959–965.
- Strominger, A. & Vafa, C. (1996). Microscopic origin of the Bekenstein-Hawking entropy. *Physics Letters B*, 379(1–4), 99–104.
- Surya, S. (2019). The causal set approach to quantum gravity. *Living Reviews in Relativity*, 22(5), 1–75. <https://arxiv.org/abs/1903.11544>.
- Susskind, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, 36(11), 6377–6396.
- Susskind, L. (2008). *The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics*. Boston: Little, Brown.
- 't Hooft, G. (1993). Dimensional reduction in quantum gravity. <https://arxiv.org/abs/gr-qc/9310026>.
- 't Hooft, G. & Veltman, M. (1974). One-loop divergencies in the theory of gravitation. *Annales de l'IHP Physique Theorique*, 20(1), 69–94.
- Thébault, K. (2019). What can we learn from analogue experiments? In R. Dardashti, R. Dawid, & K. Thébault (Eds.), *Why Trust a Theory?: Epistemology of Fundamental Physics* (pp. 184–201). Cambridge: Cambridge University Press.
- Thébault, K. (2021). The problem of time. In E. Knox & A. Wilson (Eds.), *The Routledge Companion to Philosophy of Physics* (pp. 1–15). New York: Routledge.
- Thébault, K. (2023). Big bang singularity resolution in quantum cosmology. *Classical and Quantum Gravity*, 40(5), 055007.
- Tilloy, A. (2018). Binding quantum matter and space-time, without romanticism. *Foundations of Physics*, 48, 1572–9516.
- Unruh, W. (1981). Experimental black-hole evaporation? *Physical Review Letters*, 46, 1351–1358.
- Unruh, W. G. & Wald, R. M. (2017). Information loss. *Reports on Progress in Physics*, 80(9), 092002.
- van Dongen, J., De Haro, S., Visser, M., & Butterfield, J. (2020). Emergence and correspondence for string theory black holes. *Studies in History and Philosophy of Modern Physics*, 69, 112–127.
- Verlinde, E. (2011). On the origin of gravity and the laws of Newton. *Journal of High Energy Physics*, 2011(29), 1–27.

- Verlinde, E. (2017). Emergent gravity and the dark universe. *SciPost Physics*, 2, 016.
- Visser, M., Barceló, C., & Liberati, S. (2002). Analogue models of and for gravity. *General Relativity and Gravitation*, 34(10), 1719–1734. 3rd Australasian Conference on General Relativity and Graviattion Jul 11–13, 2001 Perth, Australia.
- Volovik, G. (2003). *The Universe in a Helium Droplet*. Oxford: Oxford University Press.
- Wald, R. M. (1994). *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago: University of Chicago Press.
- Wald, R. M. (2001). The thermodynamics of black holes. *Living Reviews in Relativity*, 4(1), 6.
- Wall, A. C. (2009). Ten proofs of the generalized second law. *Journal of High Energy Physics*, 2009(06), 021.
- Wallace, D. (2006). In defence of naivete: The conceptual status of Lagrangian quantum field theory. *Synthese*, 151(1), 33.
- Wallace, D. (2011). Taking particle physics seriously: A critique of the algebraic approach to quantum field theory. *Studies in History and Philosophy of Modern Physics*, 42(2), 116–125.
- Wallace, D. (2018a). The case for black hole thermodynamics part i: Phenomenological thermodynamics. *Studies in History and Philosophy of Modern Physics*, 64, 52–67.
- Wallace, D. (2018b). The case for black hole thermodynamics part ii: Statistical mechanics. *Studies in History and Philosophy of Modern Physics*, 66(C), 103–117.
- Wallace, D. (2020). Why black hole information loss is paradoxical. In N. Huggett, K. Matsubara, & C. Wüthrich (Eds.), *Beyond Spacetime: The Foundations of Quantum Gravity* (pp. 209–236). Cambridge: Cambridge University Press.
- Wallace, D. (2022). Quantum gravity at low energies. *Studies in History and Philosophy of Science*, 94, 31–46.
- Weinberg, S. (1979). Ultraviolet divergencies in quantum theories of gravitation. In S. Hawking & W. Israel (Eds.), *General Relativity, an Einstein Centenary Survey* (pp. 790–831). Cambridge: Cambridge University Press.
- Weinberg, S. (1995). *The Quantum Theory of Fields, Vol. I (Foundations)*. Cambridge: Cambridge University Press.
- Weinberg, S. (2009). Effective field theory, past and future. *arXiv:hep-th/0908.1964v3*.
- Weinfurter, S., Tedford, E. W., Penrice, M. C., Unruh, W. G., & Lawrence, G. A. (2013). Classical aspects of Hawking radiation verified in analogue

- gravity experiment. In D. Faccio, F. Belgiorno, S. Cacciatori, V. Gorini, S. Liberati, & U. Moschella (Eds.), *Analogue Gravity Phenomenology* (pp. 167–180). Cham: Springer.
- Weinstein, S. & Rickles, D. (2021). Quantum gravity. *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/fall2021/entries/quantum-gravity/>.
- Wheeler, J. (2000). *Geons, Black Holes and Quantum Foam*. London: W.W. Norton.
- Williams, P. (2021). Renormalization group methods. In E. Knox & A. Wilson (Eds.), *The Routledge Companion to Philosophy of Physics* (pp. 1–30). New York: Routledge.
- Williams, R. (2006). Discrete quantum gravity. *Journal of Physics: Conference Series*, 33(1), 38.
- Witten, E. (1998). Anti-de Sitter space, thermal phase transition, and confinement in gauge theories. *Advances in Theoretical and Mathematical Physics*, 2, 505–532.
- Wüthrich, C. (2005). To quantize or not to quantize: Fact and folklore in quantum gravity. *Philosophy of Science*, 72, 777–788.
- Zee, A. (2010). *Quantum Field Theory in a Nutshell* (Second ed.). Princeton: Princeton University Press.
- Zeh, H. (2011). Feynmans interpretation of quantum theory. *The European Physical Journal H*, 36, 63–74.
- Zwiebach, B. (2009). *A First Course in String Theory* (Second ed.). Cambridge: Cambridge University Press.

Acknowledgements

Thanks to Jim Weatherall for letting me do this; the referees for suggestions for improvements; Niels Linnemann and Kian Salimkhani for motivation and support; Eric Curiel for comments on Section 7; Sebastiano Orioli for making (most of) the figures.

The Philosophy of Physics

James Owen Weatherall

University of California, Irvine

James Owen Weatherall is Professor of Logic and Philosophy of Science at the University of California, Irvine. He is the author, with Cailin O'Connor, of *The Misinformation Age: How False Beliefs Spread* (Yale, 2019), which was selected as a *New York Times* Editors' Choice and Recommended Reading by *Scientific American*. His previous books were *Void: The Strange Physics of Nothing* (Yale, 2016) and the *New York Times* bestseller *The Physics of Wall Street: A Brief History of Predicting the Unpredictable* (Houghton Mifflin Harcourt, 2013). He has published approximately fifty peer-reviewed research articles in journals in leading physics and philosophy of science journals and has delivered over 100 invited academic talks and public lectures.

About the Series

This Cambridge Elements series provides concise and structured introductions to all the central topics in the philosophy of physics. The Elements in the series are written by distinguished senior scholars and bright junior scholars with relevant expertise, producing balanced, comprehensive coverage of multiple perspectives in the philosophy of physics.

The Philosophy of Physics

Elements in the Series

Idealizations in Physics

Elay Shech

The Temporal Asymmetry of Causation

Alison Fernandes

Special Relativity

James Read

Philosophy of Particle Physics

Porter Williams

Foundations of Statistical Mechanics

Roman Frigg and Charlotte Werndl

From Randomness and Entropy to the Arrow of Time

Lena Zuchowski

Philosophy of Physical Magnitudes

Niels C. M. Martens

The Philosophy of Symmetry

Nicholas Joshua Yii Wye Teh

Laws of Physics

Eddy Keming Chen

Foundations of General Relativity

Samuel C. Fletcher

Gauge Theory and the Geometrization of Physics

Henrique De Andrade Gomes

Why Do We Want a Theory of Quantum Gravity?

Karen Crowther

A full series listing is available at: www.cambridge.org/EPPH